

Junhe Zhou
Meisong Tong

Optical Waves in Waveguides and Free Space

Multimode Propagation



Springer

Springer Series in Optical Sciences

Founding Editor

H. K. V. Lotsch

Volume 251

Editor-in-Chief

William T. Rhodes, Florida Atlantic University, Boca Raton, USA

Series Editors

Ali Adibi, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA

Toshimitsu Asakura, Toyohira-ku, Hokkai-Gakuen University, Sapporo, Hokkaido, Japan

Theodor W. Hänsch, Max Planck Institute of Quantum Optics, Garching b. München, Germany

Kazuya Kobayashi, Department of Electrical, Electronic, and Communication Engineering, Chuo University, Bunkyo-ku, Japan

Ferenc Krausz, Max Planck Institute of Quantum Optics, Garching b. München, Germany

Vadim Markel, Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA

Barry R. Masters, Cambridge, USA

Katsumi Midorikawa, Laser Tech Lab, RIKEN Advanced Science Institute, Saitama, Japan

Herbert Venghaus, Fraunhofer Institute for Telecommunications, Berlin, Germany

Horst Weber, Berlin, Germany

Harald Weinfurter, München, Germany

Iwao Matsuda, Institute for Solid State Physics, The University of Tokyo, Kashiwa, Chiba, Japan

Springer Series in Optical Sciences is led by Editor-in-Chief William T. Rhodes, Florida Atlantic University, USA, and provides an expanding selection of research monographs in all major areas of optics:

- lasers and quantum optics
- ultrafast phenomena
- optical spectroscopy techniques
- optoelectronics
- information optics
- applied laser technology
- industrial applications and
- synchrotron radiation and X-ray optics
- other topics of contemporary interest.

With this broad coverage of topics the series is useful to research scientists and engineers who need up-to-date reference books.

Junhe Zhou · Meisong Tong

Optical Waves in Waveguides and Free Space

Multimode Propagation

Junhe Zhou 
Department of Electronic Science
and Technology
Tongji University
Shanghai, China

Meisong Tong 
Department of Electronic Science
and Technology
Tongji University
Shanghai, China

ISSN 0342-4111 ISSN 1556-1534 (electronic)
Springer Series in Optical Sciences
ISBN 978-981-97-9546-8 ISBN 978-981-97-9547-5 (eBook)
<https://doi.org/10.1007/978-981-97-9547-5>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2024

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Preface

This book provides a thorough review of the multi-mode propagation inside the optical waveguides and free space. Multi-mode propagation in waveguides and free space has received particular attention recently due to its promising applications in various areas, such as communications and sensing. In the multi-mode waveguides and free space, the modes couple to each other due to the purposely engineered or random index fluctuations. The coupled modes interfere with one another and can be possibly exploited to form functional devices. It is revealed in this book that a universal coupled mode theory can be used to describe the mode propagation inside the waveguides and in free space. Stochastic analysis can be performed based on the theory and analytical solutions can be derived for the second-order and fourth-order moments to avoid time consuming Monte Carlo simulations. Mode generation and multiplexing devices are under consideration afterwards, which play key roles in mode division multiplexing systems. Finally, the concept is extended for the temporal domain and the four-dimensional space time.

The book provides a comprehensive review of the rapidly expanding field of multi-mode transmission system and its basic physics, covers a vast array of special topics and applications related to the core technologies of the mode division multiplexing systems, and illustrates the wide use of the multi-mode propagation and transmission systems.

The book is organized as follows.

Chapter 1 “Multimode Propagation Fundamentals”. The chapter reviews the fundamental physics for multi-mode propagation inside waveguides and in free space. Coupled mode theory is derived. Multi-mode interference theory is revealed.

Chapter 2 “Multimode Interference in Waveguides and Free Space”. The chapter deals with the modes in the waveguides or free space which propagates with different velocities and propagation constants. Interference is expected and can be used to design useful functional devices such as multi-mode interference couplers. Functional devices such as unitary transformers are designed based on the multi-mode interference concept. In free space, the multi-mode interference is described as the Talbot effect, which can be exploited to generate spatial modes.

Chapter 3 “Coupled Mode Theory for Multi-mode Propagation”. The chapter introduces the coupled mode theory, which is a powerful tool for multi-mode propagation analysis. The coupled mode theory will be formulated both for the classic optics and quantum optics. While the coupled mode theory is a powerful tool for multi-mode propagation analysis in classic optics, quantum formulation of the multimode propagation is provided afterwards.

Chapter 4 “Wave Propagation in Multi-core Fibers”. The chapter studies the multi-core fibers, which are considered as a special kind of multi-mode waveguide while the cores are closely coupled to each other and form the super-modes. The concept is discussed and analytical solutions are given for various multi-core fibers with different core arrangement.

Chapter 5 “Random Coupling Theory for Multi-mode Waveguides and Free Space Propagation”. The chapter studies the random coupled mode theory, which has the coupling coefficients varies randomly and should be studied through the stochastic approach. The basic equations are given and the analytical results for the average value, as well as the second- and fourth-order moments are studied analytically as well.

Chapter 6 “Mode Generation and Multiplexing for Multi-mode Waveguides and Free Space”. The chapter studies the mode generator and multiplexer, which are the key components for mode division multiplexing systems. The chapter discusses the novel ways to realize mode generators and multiplexers.

Chapter 7 “Temporal Modes and Spatiotemporal Modes”. The chapter extends the theory of spatial beam propagation to the temporal domain, and the concept of temporal waveguides and temporal fractional Fourier transform are proposed. Afterwards, a temporal signal processing tool, i.e., a temporal diffractive deep neural network is proposed. The concept is extended beyond the spatial domain or the temporal domain, but the four-dimensional space time. A new class of spatial-temporal beam is discussed and a spatial-temporal diffractive deep neural network is proposed in the end. Finally, the authors would like to acknowledge the proofreading efforts of associate professor Jing Zheng from Shanghai Institute of Technology, as well as the contributions of the PhD students Tengyuan Liu and Chen Chen from Tongji University.

Shanghai, China

Junhe Zhou
Meisong Tong

Contents

1	Multimode Propagation Fundamentals	1
1.1	Maxwell Equation	1
1.2	Mode Orthogonal Relation	4
1.3	Vectorial Coupled Mode Theory	5
1.4	Simple Planar Metal Waveguide and Multimode Interference	9
	References	11
2	Multimode Interference in Waveguides and Free Space	13
2.1	All Optical Discrete Fourier Transform Based on Multimode Interference Couplers	13
2.1.1	Introduction	14
2.1.2	Theory	14
2.1.3	Simulation Results and Discussion	18
2.1.4	Summary	20
2.2	Arbitrary Periodical Pattern Generation Based on Fractional Talbot Effect	20
2.2.1	Introduction	20
2.2.2	Theory	21
2.2.3	Numerical Results with Infinite Periodical Patterns	26
2.2.4	Experimental Verifications	28
2.2.5	Summary	30
2.3	Ring Shape Talbot Effect	32
2.3.1	Introduction	32
2.3.2	Theory	33
2.3.3	Results and Discussions	35
2.3.4	Summary	40
	References	41
3	Coupled Mode Theory for Multi-mode Propagation	45
3.1	Coupled Mode Theory in Classical Optics	45
3.1.1	<i>Coupled Mode Theory Inside Optical Waveguides with the Scalar Wave Equation</i>	45

3.1.2	Extension of Formulas in the Vector Theory	47
3.2	Characterization of PT-Symmetric Quantum Interference	
	Based on the Coupled Mode Theory	48
3.2.1	Introduction	48
3.2.2	Coupled Mode Theory for the Coupled Waveguides	
	with Unequal Losses	49
3.2.3	Analysis of PT Symmetric Quantum Entanglement	51
3.2.4	Summary	57
3.3	Analytical Formulation of Quantum Interference Inside	
	Coupled Waveguides with Unequal Losses	57
3.3.1	Introduction	57
3.3.2	Coupled Mode Theory for the Coupled Waveguides	
	with Unequal Losses	59
3.3.3	Analysis of Anti-PT Symmetric Quantum	
	Entanglement in a Three-Waveguide System	61
3.3.4	Analysis of the Quantum State Evolution	
	in an 89-Waveguide System	67
3.3.5	Summary	69
3.3.6	Proof of the Preservation of the Commutation Property	70
3.3.7	Proof of Eq. (3.37)	71
3.3.8	Derivation of the Diagonal Terms of the Density	
	Matrix with Input as $ 1, 0, 1\rangle$	73
	References	74
4	Wave Propagation in Multi-core Fibers	77
4.1	Analytical Formulation of Super-Modes Inside Multi-core	
	Fibers with Circularly Distributed Cores	77
4.1.1	Introduction	78
4.1.2	MCFs with Cores in One Ring	79
4.1.3	MCFs with Cores in Two Rings	81
4.1.4	MCFs with Cores in Three Rings	87
4.1.5	MCFs with Cores in L Rings	89
4.1.6	MCFs with a Core in the Middle of the Rings	90
4.1.7	Case of Strong Coupling Among the Cores	94
4.1.8	Summary	96
4.2	A Non-orthogonal Coupled Mode Theory for Super-Modes	
	Inside Multi-core Fibers	97
4.2.1	Introduction	97
4.2.2	A Rigorous Non-Orthogonal Coupled Mode Theory	98
4.2.3	Analytical Formulation of Super-Modes Inside	
	Circularly Aligned MCFs	99
4.2.4	Comparison with the Published Results	101
4.2.5	Numerical Simulations and Discussions	102
4.2.6	Signal Pre-Coding Based on the Concept	103
4.2.7	De-Multiplexing the Signals Based on IDFT	106

4.2.8	Summary	107
4.3	Analytical Study of Super-Modes in Multi-core Fibers with Multimode Cores	108
4.3.1	Introduction	108
4.3.2	Rigorous Non-Orthogonal Coupled Mode Equations for Super-Modes Characterization in Circular MCFs	109
4.3.3	Super-Modes Formed by Non-Degenerated Modes	112
4.3.4	Super-Modes Formed by Two Degenerated Modes	113
4.3.5	Super-Modes Formed by Three Degenerated Modes	115
4.3.6	Super-Modes Formed by L Degenerated Modes	117
4.3.7	Numerical Simulations and Discussions	118
4.3.8	Summary	120
4.4	Power Splitting and Switching in a Multi-core Fiber Based on the Multimode Interference Effect	121
4.4.1	Introduction	122
4.4.2	Theory for the MMI Effect Inside a Ring Core Fiber	124
4.4.3	Working Principle of the Ring Shape MMI Coupler Based Power Splitter and Switch	126
4.4.4	Numerical Simulations	127
4.4.5	Summary	131
	References	132
5	Random Coupling Theory for Multi-mode Waveguides and Free Space Propagation	137
5.1	A Comprehensive Analytical Model to Characterize Randomness in Optical Waveguides	137
5.1.1	Introduction	138
5.1.2	Mathematical Model	139
5.1.3	Simulation Results and Discussion	149
5.1.4	Summary	157
5.2	Analytical Study on the Evolutionary Behavior of Transfer Matrix Element Moments in Strongly Coupled Multimode Systems	158
5.2.1	Introduction	158
5.2.2	Theory	160
5.2.3	Asymptotic Behaviors of the Matrix Element Moments	166
5.2.4	Results and Discussions	168
5.2.5	Summary	172
5.3	Analytical Expressions for the Crosstalk of Super-Modes in the Tightly Bounded Multicore Fibers	174
5.3.1	Introduction	175
5.3.2	Theory	177
5.3.3	Results and Discussions	181

5.3.4	Summary	187
5.4	On the Variance of Crosstalk in Multi-core Fibers: With and Without the Polarization Effect	188
5.4.1	Introduction	188
5.4.2	Theory	189
5.4.3	Results and Discussions	194
5.4.4	Summary	197
5.5	A Coupled Mode Theory for Orbital Angular Momentum Modes Transmission in the Presence of Atmosphere Turbulence	197
5.5.1	Introduction	198
5.5.2	The Coupled Power Equations	199
5.5.3	Results and Discussions	205
5.5.4	Summary	213
5.6	Capacity Analysis for Free Space Coherent Optical MIMO Transmission Systems: With and Without Adaptive Optics	214
5.6.1	Introduction	214
5.6.2	Theory	216
5.6.3	Results and Discussions	220
5.6.4	Summary	226
	Appendix	227
	References	230
6	Mode Generation and Multiplexing for Multi-mode Waveguides and Free Space	237
6.1	A Novel Mode Multiplexer/De-Multiplexer for Multi-core Fibers	237
6.1.1	Introduction	238
6.1.2	Mathematical Description	239
6.1.3	Numerical Results and Discussion	242
6.1.4	Summary	245
6.2	OAM States Generation/Detection Based on the Multimode Interference Effect in a Ring Core Fiber	246
6.2.1	Introduction	247
6.2.2	MMI Theory in Ring Core Fibers	248
6.2.3	MMI Effect Based OAM MUX and DEMUX	252
6.2.4	Numerical Simulations	252
6.2.5	Summary	260
6.3	A Tunable Arbitrary Unitary Transformer Based on Multiple Sections of Multicore Fibers with Phase Control	261
6.3.1	Introduction	261
6.3.2	Theory	262
6.3.3	Results and Discussions	265
6.3.4	Summary	273

6.3.5	The Proof that an Arbitrary Unitary Transformer Can Be Realized Based on the Proposed Structure	274
6.3.6	Realization of a 3×3 Arbitrary Unitary Transformer	277
6.4	Orbital Angular Momentum Mode Sorting Based on a Hybrid Radial-Angular Hybrid Lens	279
6.4.1	Introduction	280
6.4.2	Theory	281
6.4.3	Results and Discussions	284
6.4.4	Summary	288
6.4.5	The Detailed Derivation for the Angular Lens Equation with the Specific Input Ring and Output Ring Radii	288
	References	292
7	Temporal Modes and Spatiotemporal Modes	297
7.1	Comprehensive Study on the Concept of Temporal Optical Waveguides	297
7.1.1	Introduction	298
7.1.2	Theory	299
7.1.3	Numerical Results and Discussions	302
7.1.4	Summary	307
7.2	Fractional Fourier Transform Based on a Temporal Gradient-Index Lens	307
7.2.1	Introduction	307
7.2.2	Theory	308
7.2.3	Results and Discussions	310
7.2.4	Summary	314
7.3	Nonlinear Fourier Transform Receiver Based on a Time Domain Diffractive Deep Neural Network	314
7.3.1	Introduction	315
7.3.2	The Time Domain D2NN and All Optical Error Back-Propagation	316
7.3.3	Results and Discussions	319
7.3.4	Summary	326
7.4	Spherical Gauss-Laguerre Beam Propagation in 4D Space-time	327
7.4.1	Introduction	327
7.4.2	Theory	328
7.4.3	Results and Discussions	332
7.4.4	Summary	336
7.5	Spatiotemporal Diffractive Deep Neural Networks	337
7.5.1	Introduction	337
7.5.2	The STD2NN Architecture and All Optical Error Back-Propagation	339
7.5.3	Results and Discussions	345

7.5.4 Summary	352
References	352
Index	357

Chapter 1

Multimode Propagation Fundamentals



Abstract The chapter reviews the fundamental physics for multi-mode propagation inside waveguides and in free space. Coupled mode theory is derived. Multimode interference theory is revealed.

This chapter offers an overall review on the fundamental physics for multi-mode propagation inside waveguides and in free space. The fundamental equations for mode propagation are derived. Most of the results have been presented by the pioneering researchers [1–3], which are presented in this chapter once again for readers' convenience. The concept of multimode interference for the waveguides is discussed at the end of the chapter.

1.1 Maxwell Equation

The fundamental equation for the electromagnetic field is [1]:

$$\begin{aligned}\nabla \times \mathbf{H} &= \mathbf{J} + \frac{\partial \mathbf{D}}{\partial t}, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{D} &= \rho,\end{aligned}\tag{1.1}$$

where $\mathbf{E}$ is the electric field, $\mathbf{H}$ is the magnetic field, $\mathbf{D}$ is the electric flux, $\mathbf{B}$ is the magnetic flux, $\mathbf{J}$ is the current density, and ρ is the charge density. $\mathbf{E}$ and $\mathbf{D}$, as well as $\mathbf{B}$ and $\mathbf{H}$, are related to each other respectively [1]:

$$\begin{aligned}\mathbf{D} &= \varepsilon \mathbf{E} = \varepsilon_r \varepsilon_0 \mathbf{E}, \\ \mathbf{B} &= \mu \mathbf{H} = \mu_r \mu_0 \mathbf{H},\end{aligned}\tag{1.2}$$

where ε is the permittivity, μ is the permeability, ε_0 is the permittivity in free space, μ_0 is the permeability in free space, ε_r is the relative permittivity and μ_r is the relative permeability. During the study of wave propagation, the source is not under consideration. Therefore, the wave propagation in the media is described by [1]:

$$\begin{aligned}\nabla \times \mathbf{H} &= \frac{\partial \mathbf{D}}{\partial t}, \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t}, \\ \nabla \cdot \mathbf{B} &= 0, \\ \nabla \cdot \mathbf{D} &= 0.\end{aligned}\tag{1.3}$$

From the engineering point of view, the field can be studied in the frequency domain, i.e., the harmonic field can be assumed with the angular frequency of ω and the temporal harmonic factor is:

$$e^{j\omega t}.\tag{1.4}$$

It should be noted that factor can be also chosen as $e^{-j\omega t}$ without loss of generality. Therefore, we have:

$$\begin{aligned}\nabla \times \mathbf{H} &= j\omega\varepsilon\mathbf{E}, \\ \nabla \times \mathbf{E} &= -j\omega\mu_0\mathbf{H}.\end{aligned}\tag{1.5}$$

The two divergence equations in (1.4) can be omitted, which can be obtained by taking the divergence on both sides of the equation. To study the wave propagation in the waveguides and free space, we assume the propagation direction is z and the Nabla operator can be decomposed as [1]:

$$\nabla = \nabla_t + \frac{\partial}{\partial z}e_z,\tag{1.6}$$

where the lower-case t stands for the transverse and e_z is the unit vector in the z direction. By decomposing Eq. (1.5) in the transverse direction and the z direction, we have [1]:

$$\begin{aligned}\nabla_t \times (H_z e_z) + e_z \frac{\partial}{\partial z} \times \mathbf{H}_t &= j\omega\varepsilon\mathbf{E}_t, \\ \nabla_t \times \mathbf{H}_t &= j\omega\varepsilon\mathbf{E}_z, \\ \nabla_t \times (E_z e_z) + e_z \frac{\partial}{\partial z} \times \mathbf{E}_t &= -j\omega\mu_0\mathbf{H}_t, \\ \nabla_t \times \mathbf{E}_t &= -j\omega\mu_0\mathbf{H}_z,\end{aligned}\tag{1.7}$$

which suggests the z direction components can be represented by the transverse components [1]:

$$\begin{aligned}\mathbf{E}_z &= \frac{1}{j\omega\epsilon} \nabla_t \times \mathbf{H}_t, \\ \mathbf{H}_z &= -\frac{1}{j\omega\mu_0} \nabla_t \times \mathbf{E}_t.\end{aligned}\tag{1.8}$$

Substituting the first equation in Eq. (1.7) into the third equation in Eq. (1.7), we have:

$$j\omega\epsilon \nabla_t \times (E_z e_z) + e_z \frac{\partial}{\partial z} \times \left(\left(\nabla_t \times (H_z e_z) + e_z \frac{\partial}{\partial z} \times \mathbf{H}_t \right) \right) = \omega^2 \mu_0 \epsilon \mathbf{H}_t. \tag{1.9}$$

In the waveguide structures, the z dependent factor is a pure phase factor for the mode:

$$e^{-j\beta z}, \tag{1.10}$$

where β is the propagation constant of the mode. It should be noted that the sign of the phase factor should be negative with respect to the temporal harmonic factor. In this case, the partial derivative with respect to z can be replaced by $-j\beta$ [1].

$$j\omega\epsilon \nabla_t \times (E_z e_z) - j\beta e_z \times \nabla_t \times (H_z e_z) - \beta^2 e_z \times (e_z \times \mathbf{H}_t) = \omega^2 \mu_0 \epsilon \mathbf{H}_t. \tag{1.11}$$

Using the property [1]:

$$\begin{aligned}e_z \times \nabla_t \times (H_z e_z) &= \nabla_t H_z, \\ e_z \times (e_z \times \mathbf{H}_t) &= -\mathbf{H}_t,\end{aligned}\tag{1.12}$$

we have:

$$j\omega\epsilon \nabla_t \times (E_z e_z) - j\beta \nabla_t H_z + \beta^2 \mathbf{H}_t = \omega^2 \mu_0 \epsilon \mathbf{H}_t. \tag{1.13}$$

This suggest that the transverse field components of the modes can be represented by their z direction field components [1]:

$$\begin{aligned}\mathbf{H}_t &= \frac{1}{\omega^2 \mu_0 \epsilon - \beta^2} (j\omega\epsilon \nabla_t \times (E_z e_z) - j\beta \nabla_t H_z), \\ \mathbf{E}_t &= \frac{1}{\omega^2 \mu_0 \epsilon - \beta^2} (-j\omega\mu_0 \nabla_t \times (H_z e_z) - j\beta \nabla_t E_z).\end{aligned}\tag{1.14}$$

In summary, the transverse direction and the z direction field components can be related to each other [1]:

$$\begin{aligned}
\mathbf{E}_z &= \frac{1}{j\omega\varepsilon} \nabla_t \times \mathbf{H}_t, \\
\mathbf{H}_z &= -\frac{1}{j\omega\mu_0} \nabla_t \times \mathbf{E}_t, \\
\mathbf{H}_t &= \frac{1}{\omega^2\mu_0\varepsilon - \beta^2} (j\omega\varepsilon \nabla_t \times (E_z \mathbf{e}_z) - j\beta \nabla_t H_z), \\
\mathbf{E}_t &= \frac{1}{\omega^2\mu_0\varepsilon - \beta^2} (-j\omega\mu_0 \nabla_t \times (H_z \mathbf{e}_z) - j\beta \nabla_t E_z).
\end{aligned} \tag{1.15}$$

1.2 Mode Orthogonal Relation

Assuming the waveguide structure have two modes with different field distributions and different propagation constants with the upper-case letter m and n indicating the mode order, we may prove that two modes are orthogonal.

According to Eq. (1.5), we have [1]:

$$\begin{aligned}
\nabla \times \mathbf{H}^n &= j\omega\varepsilon \mathbf{E}^n, \\
\nabla \times \mathbf{E}^n &= -j\omega\mu_0 \mathbf{H}^n, \\
\nabla \times \mathbf{H}^m &= j\omega\varepsilon \mathbf{E}^m, \\
\nabla \times \mathbf{E}^m &= -j\omega\mu_0 \mathbf{H}^m.
\end{aligned} \tag{1.16}$$

It is straight forward to have [1]:

$$\begin{aligned}
(\nabla \times \mathbf{H}^{n*}) \cdot \mathbf{E}^m &= -j\omega\varepsilon \mathbf{E}^{n*} \cdot \mathbf{E}^m, \\
(\nabla \times \mathbf{E}^{n*}) \cdot \mathbf{H}^m &= j\omega\mu_0 \mathbf{H}^{n*} \cdot \mathbf{H}^m, \\
(\nabla \times \mathbf{H}^m) \cdot \mathbf{E}^{n*} &= j\omega\varepsilon \mathbf{E}^m \cdot \mathbf{E}^{n*}, \\
(\nabla \times \mathbf{E}^m) \cdot \mathbf{H}^{n*} &= -j\omega\mu_0 \mathbf{H}^m \cdot \mathbf{H}^{n*}.
\end{aligned} \tag{1.17}$$

Adding the first equation with the third equation and adding the second equation with the fourth equation in Eq. (1.17), we have [1]:

$$\begin{aligned}
(\nabla \times \mathbf{H}^{n*}) \cdot \mathbf{E}^m + (\nabla \times \mathbf{H}^m) \cdot \mathbf{E}^{n*} &= 0, \\
(\nabla \times \mathbf{E}^{n*}) \cdot \mathbf{H}^m + (\nabla \times \mathbf{E}^m) \cdot \mathbf{H}^{n*} &= 0.
\end{aligned} \tag{1.18}$$

Combining the two equations, we have [1]:

$$\nabla \cdot (\mathbf{E}^m \times \mathbf{H}^{n*} + \mathbf{E}^{n*} \times \mathbf{H}^m) = 0. \tag{1.19}$$

Decomposing the Nabla operator and the fields into the transverse components and the z components, we have [1]:

$$\nabla_t \cdot (\mathbf{E}^m \times \mathbf{H}^{n*} + \mathbf{E}^{n*} \times \mathbf{H}^m) + \frac{\partial}{\partial z} e_z \cdot (\mathbf{E}_t^m \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^m) = 0. \quad (1.20)$$

Replacing the partial derivative with respect to z with the propagation constants for mode m and mode n, we have [1]:

$$\nabla_t \cdot (\mathbf{E}^m \times \mathbf{H}^{n*} + \mathbf{E}^{n*} \times \mathbf{H}^m) + e_z \cdot (\beta_m - \beta_n)(\mathbf{E}_t^m \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^m) = 0. \quad (1.21)$$

Taking integration on the transverse plane, we have [1]:

$$\begin{aligned} & \iint \nabla_t \cdot (\mathbf{E}^m \times \mathbf{H}^{n*} + \mathbf{E}^{n*} \times \mathbf{H}^m) dS \\ & + \iint e_z \cdot (\beta_m - \beta_n)(\mathbf{E}_t^m \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^m) dS = 0. \end{aligned} \quad (1.22)$$

The first integral vanishes as the integration goes toward infinity, and we have [1]:

$$(\beta_m - \beta_n) \iint e_z \cdot (\mathbf{E}_t^m \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^m) dS = 0. \quad (1.23)$$

If the nth mode propagates backwardly, we have [1]:

$$(\beta_m + \beta_n) \iint e_z \cdot (\mathbf{E}_t^m \times \mathbf{H}_t^{n*} - \mathbf{E}_t^{n*} \times \mathbf{H}_t^m) dS = 0. \quad (1.24)$$

Therefore, we have [1]:

$$\iint e_z \cdot (\mathbf{E}_t^m \times \mathbf{H}_t^{n*}) dS = 0, \quad (1.25)$$

which is the general orthogonal relation between the two modes.

1.3 Vectorial Coupled Mode Theory

In this section, the mode coupling relation in the waveguide structure is derived based on the vectorial formation. The scalar formulation is elaborated in the subsequent chapter and will not be discussed in this section.

As is discussed above, the mode n in the unperturbed waveguide structure fulfills [2]:

$$\begin{aligned}
& \frac{1}{j\omega} \nabla_t \times \left(\frac{1}{\varepsilon} \nabla_t \times \mathbf{H}_t^n \right) - j\beta_n e_z \times \mathbf{E}_t^n = -j\omega\mu_0 \mathbf{H}_t^n, \\
& -\frac{1}{j\omega\mu_0} \nabla_t \times (\nabla_t \times \mathbf{E}_t^n) - j\beta_n e_z \times \mathbf{H}_t^n = j\omega\varepsilon \mathbf{E}_t^n.
\end{aligned} \tag{1.26}$$

The permittivity disturbance results in the other field distribution as [2]:

$$\begin{aligned}
& \frac{1}{j\omega} \nabla_t \times \left(\frac{1}{\varepsilon'} \nabla_t \times \mathbf{H}_t \right) + e_z \frac{\partial}{\partial z} \times \mathbf{E}_t = -j\omega\mu_0 \mathbf{H}_t, \\
& -\frac{1}{j\omega\mu_0} \nabla_t \times (\nabla_t \times \mathbf{E}_t) + e_z \frac{\partial}{\partial z} \times \mathbf{H}_t = j\omega\varepsilon' \mathbf{E}_t.
\end{aligned} \tag{1.27}$$

Since the modes have been proved to be the orthonormal basis, the field with perturbed permittivity can be expressed by the modes of the unperturbed waveguide structure [2]:

$$\begin{aligned}
\mathbf{E}_t &= \sum a_n \mathbf{E}_t^n, \\
\mathbf{H}_t &= \sum b_n \mathbf{H}_t^n,
\end{aligned} \tag{1.28}$$

where a_n and b_n are coefficients for the mode expansion series. Substituting Eq. (1.28) into Eq. (1.27), we have:

$$\begin{aligned}
& \frac{1}{j\omega} \sum b_n \nabla_t \times \left(\frac{1}{\varepsilon'} \nabla_t \times \mathbf{H}_t^n \right) + \sum \frac{\partial}{\partial z} a_n (e_z \times \mathbf{E}_t^n) - j\beta_n \sum a_n (e_z \times \mathbf{E}_t^n) \\
& = -j\omega\mu_0 \sum b_n \mathbf{H}_t^n, \\
& -\frac{1}{j\omega\mu_0} \sum a_n \nabla_t \times (\nabla_t \times \mathbf{E}_t^n) + \sum \frac{\partial}{\partial z} b_n (e_z \times \mathbf{H}_t^n) - j\beta_n \sum b_n (e_z \times \mathbf{H}_t^n) \\
& = j\omega\varepsilon' \sum a_n \mathbf{E}_t^n.
\end{aligned} \tag{1.29}$$

Using Eq. (1.26), we have:

$$\begin{aligned}
& \sum \left(\frac{\partial b_n}{\partial z} - j\beta_n b_n + j\beta_n a_n \right) (e_z \times \mathbf{H}_t^n) = j\omega(\varepsilon' - \varepsilon) \sum a_n \mathbf{E}_t^n, \\
& \sum \left(\frac{\partial a_n}{\partial z} - j\beta_n a_n + j\beta_n b_n \right) (e_z \times \mathbf{E}_t^n) = -\frac{1}{j\omega} \sum b_n \nabla_t \times \left(\left(\frac{1}{\varepsilon'} - \frac{1}{\varepsilon} \right) \nabla_t \times \mathbf{H}_t^n \right).
\end{aligned} \tag{1.30}$$

By assuming:

$$\begin{aligned} a'_n &= a_n e^{-j\beta_n z}, \\ b'_n &= b_n e^{-j\beta_n z}, \end{aligned} \quad (1.31)$$

We have [2]:

$$\begin{aligned} \sum \left(\frac{\partial b_n}{\partial z} + j\beta_n a_n \right) (e_z \times \mathbf{H}_t^n) &= j\omega(\varepsilon' - \varepsilon) \sum a_n \mathbf{E}_t^n, \\ \sum \left(\frac{\partial a_n}{\partial z} + j\beta_n b_n \right) (e_z \times \mathbf{E}_t^n) &= -\frac{1}{j\omega} \sum b_n \nabla_t \times \left(\left(\frac{1}{\varepsilon'} - \frac{1}{\varepsilon} \right) \nabla_t \times \mathbf{H}_t^n \right). \end{aligned} \quad (1.32)$$

For simplicity, we have dropped the upper-case prime in Eq. (1.32). Taking the integration on the transverse plane and using the orthogonal relation derived in the previous section, we have [2]:

$$\begin{aligned} \frac{\partial b_n}{\partial z} + j\beta_n a_n &= \sum K_{mn} a_m, \\ \frac{\partial a_n}{\partial z} + j\beta_n b_n &= \sum k_{mn} b_m, \end{aligned} \quad (1.33)$$

where [2]:

$$\begin{aligned} K_{mn} &= -\frac{j\omega \iint (\varepsilon' - \varepsilon) \mathbf{E}_t^{n*} \cdot \mathbf{E}_t^m dS}{\iint e_z \cdot (\mathbf{E}_t^{n*} \times \mathbf{H}_t^n) dS}, \\ k_{mn} &= -\frac{1}{j\omega} \frac{\iint \nabla_t \times \left(\left(\frac{1}{\varepsilon'} - \frac{1}{\varepsilon} \right) \nabla_t \times \mathbf{H}_t^m \right) \cdot \mathbf{H}_t^{n*} dS}{\iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*}) dS}. \end{aligned} \quad (1.34)$$

The second class of coupling coefficients k_{mn} can be simplified after some calculations [2]:

$$k_{mn} = -\frac{j\omega \iint \frac{\varepsilon}{\varepsilon'} (\varepsilon' - \varepsilon) \mathbf{E}_z^{n*} \cdot \mathbf{E}_z^m dS}{\iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*}) dS}. \quad (1.35)$$

Considering the solution to Eq. (1.33) when the coupling coefficients are zero [2]:

$$\begin{aligned} \frac{\partial b_n}{\partial z} + j\beta_n a_n &= 0, \\ \frac{\partial a_n}{\partial z} + j\beta_n b_n &= 0, \end{aligned} \quad (1.36)$$

we have the solution as [2]:

$$\begin{aligned} a_n &= c_n^+ \exp(-j\beta_n z) + c_n^- \exp(j\beta_n z), \\ b_n &= c_n^+ \exp(-j\beta_n z) - c_n^- \exp(j\beta_n z), \end{aligned} \quad (1.37)$$

where the sign $+$ indicates the forward propagating waves and $-$ the backward propagating waves, and c is the wave amplitude. It can be rewritten as [2]:

$$\begin{aligned} a_n &= a_n^+ + a_n^-, \\ b_n &= a_n^+ - a_n^-. \end{aligned} \quad (1.38)$$

Substituting Eq. (1.38) into Eq. (1.33), we have [2]:

$$\begin{aligned} \frac{da_n^+}{dz} &= -j\beta_n a_n^+ + \sum (K_{mn}^{+,+} a_m^+ + K_{mn}^{+,-} a_m^-), \\ \frac{da_n^-}{dz} &= j\beta_n a_n^+ + \sum (K_{mn}^{-,+} a_m^+ + K_{mn}^{-,-} a_m^-), \end{aligned} \quad (1.39)$$

where [2]:

$$K_{mn}^{p,q} = \frac{1}{2}(pK_{mn} + qk_{mn}), \quad (1.40)$$

where p and q are the position and the negative signs. For forward propagating modes, the coupled mode equation becomes:

$$\frac{da_n^+}{dz} = -j\beta_n a_n^+ + \sum K_{mn}^{+,+} a_m^+, \quad (1.41)$$

with the coupling coefficient as:

$$K_{mn}^{+,+} = \frac{1}{2}(K_{mn} + k_{mn}). \quad (1.42)$$

It is shown in [2] that:

$$\iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*}) dS = 2P_n \frac{\beta_n^*}{|\beta_n|}, \quad (1.43)$$

where P_n is the mode power. For the propagating modes, β_n is a real constant, and therefore, we have:

$$\begin{aligned} \iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*}) dS &= \iint e_z \cdot (\mathbf{E}_t^{n*} \times \mathbf{H}_t^n) dS = 2P_n \\ \iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^n) dS &= 4P_n \end{aligned} \quad (1.44)$$

Therefore, Eq. (1.42) can be reformulated as:

$$K_{mn}^{+,+} = -\frac{j\omega \iint (\varepsilon' \varepsilon) (\mathbf{E}_t^{n*} \cdot \mathbf{E}_t^m + \frac{\varepsilon}{\varepsilon'} \mathbf{E}_z^{n*} \cdot \mathbf{E}_z^m) dS}{\iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^n) dS}. \quad (1.45)$$

For the weak perturbation, with ε/ε' close to 1, we have:

$$K_{mn}^{+,+} = -\frac{j\omega \iint (\varepsilon' - \varepsilon)(\mathbf{E}^{n*} \cdot \mathbf{E}^m)dS}{\iint e_z \cdot (\mathbf{E}_t^n \times \mathbf{H}_t^{n*} + \mathbf{E}_t^{n*} \times \mathbf{H}_t^n)dS}. \quad (1.46)$$

1.4 Simple Planner Metal Waveguide and Multimode Interference

In this section, the physics behind the multimode waveguide is discussed and the modes for the planner waveguides will be presented. The concept is extended from 1 to 2D, i.e., the rectangular waveguide at the end of the section.

Instead of choosing the planner dielectric waveguide [4], which is often used for the illustration of the multimode interference effect, the section uses the planner metal waveguide. The conclusion can be simply extended for the dielectric waveguide by considering the Goos-Hanchen shift when evaluating the width of the waveguide.

In this planner waveguide, the y direction is considered to be uniform and the related derivative can be regarded as 0.

To start with, with the partial derivative with respect to y as 0, we express Eq. (1.5) as [3]:

$$\begin{aligned} j\omega\varepsilon_0 E_x &= -\frac{\partial H_y}{\partial z}, \\ j\omega\varepsilon_0 E_y &= \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \\ j\omega\varepsilon_0 E_z &= \frac{\partial H_y}{\partial x}, \\ -j\omega\mu_0 H_x &= -\frac{\partial E_y}{\partial z}, \\ -j\omega\mu_0 H_y &= \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}, \\ -j\omega\mu_0 H_z &= \frac{\partial E_y}{\partial x}. \end{aligned} \quad (1.47)$$

Obviously, the wave can be classified as two independent categories, i.e., the TM modes and the TE modes [3]:

$$\begin{aligned} j\omega\varepsilon_0 E_x &= -\frac{\partial H_y}{\partial z}, \\ j\omega\varepsilon_0 E_z &= \frac{\partial H_y}{\partial x}, \\ -j\omega\mu_0 H_y &= \frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x}. \end{aligned} \quad (1.48)$$

$$\begin{aligned}
j\omega\epsilon_0 E_y &= \frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x}, \\
-j\omega\mu_0 H_x &= -\frac{\partial E_y}{\partial z}, \\
-j\omega\mu_0 H_z &= \frac{\partial E_y}{\partial x}.
\end{aligned} \tag{1.49}$$

Without loss of generality, we consider the TE modes and obtain [3]

$$\begin{aligned}
\frac{\partial^2 E_y}{\partial x^2} + \frac{\partial^2 E_y}{\partial z^2} + k_0^2 E_y &= 0, \\
k_0^2 &= \omega^2 \mu_0 \epsilon_0.
\end{aligned} \tag{1.50}$$

Consider the z dependent term $e^{-j\beta z}$, we have [3]:

$$\begin{aligned}
E_y &= U e^{-j\beta z}, \\
\frac{\partial^2 U}{\partial x^2} + k_x^2 U &= 0, \\
k_x^2 &= k_0^2 - \beta^2.
\end{aligned} \tag{1.51}$$

Assuming the metal waveguide has the width of W , it can be solved by considering the two boundary conditions:

$$\begin{aligned}
U_m &= \sin\left(\frac{m\pi}{W}x\right) \quad (0 < x < W), \\
E_y^m &= \sin\left(\frac{m\pi}{W}x\right) e^{-j\sqrt{k_0^2 - \left(\frac{m\pi}{W}\right)^2}z} \quad (0 < x < W),
\end{aligned} \tag{1.52}$$

where m is the mode number, which is an integer. Hence, we have an arbitrary beam within the metal waveguide as:

$$E_y = \sum a_m E_y^m = \sum a_m \sin\left(\frac{m\pi}{W}x\right) e^{-j\sqrt{k_0^2 - \left(\frac{m\pi}{W}\right)^2}z}. \tag{1.53}$$

For the paraxial propagation, we have

$$\begin{aligned}
&\sqrt{k_0^2 - \left(\frac{m\pi}{W}\right)^2} \\
&\approx k_0 \left(1 - \frac{1}{2} \left(\frac{m\pi}{k_0 W}\right)^2\right) \\
&= k_0 - \frac{1}{2k_0} \left(\frac{m\pi}{W}\right)^2.
\end{aligned} \tag{1.54}$$

The first term is a common phase, which can be ignored during the analysis. For the second term, when:

$$z = \frac{2k_0 W^2}{\pi} = \frac{4W^2}{\lambda_0}, \quad (1.55)$$

we have:

$$\frac{1}{2k_0} \left(\frac{\pi}{W} \right)^2 z = \pi. \quad (1.56)$$

The input field E_y will reappear at the position of z with an inverse distribution in the x direction, which is referred to as the multimode interference. The effect can take place at the multiple of the above distance z . It can also take place at the fractional of z as shown in the latter chapters.

The above-mentioned analysis can be readily extended for the 2D case with a square shape hollow waveguide and the distance for the multimode interference to take place is still Eq. (1.55).

References

1. W.C. Chew, *Lectures on Theory of Microwave and Optical Waveguides*. [arXiv:2107.09672](https://arxiv.org/abs/2107.09672) (2021)
2. D. Marcuse, *Theory of dielectric optical waveguides* (Academic Press, 1991)
3. K. Kawano, T. Kitoh, *Introduction to Optical Waveguide Analysis: Solving Maxwell's Equations and the Schrödinger Equation* (Wiley, 2001)
4. L.B. Soldano, E.C.M. Pennings, Optical multimode interference devices based on self-imaging—principles and applications. *J. Lightwave Technol.* **13**(4), 615–627 (1995)

Chapter 2

Multimode Interference in Waveguides and Free Space



Abstract Multi-mode interference in waveguides and free space. The chapter deals with the modes in the waveguides or free space which propagates with different velocities and propagation constants. Interference is expected and can be used to design useful functional devices such as multi-mode interference couplers. Functional devices such as unitary transformers are designed based on the multi-mode interference concept. In free space, the multi-mode interference is described as the Talbot effect, which can be exploited to generate spatial modes.

As is shown in the previous chapter, multimode waveguide could possess multiple modes with different propagation constants. If the difference of the propagation phases equals the multiple of 2π , the field formed by the different modes could reappear at certain propagation distance and this is called multimode interference. Multimode interference can not only take place in the above-mentioned positions, but also in the positions with the propagation phases equaling the fraction of 2π , which is referred to as the fractional multimode interference. Furthermore, the multimode interference effect can occur in free space and is described as the spatial Talbot effect, which can be exploited to generate spatial modes. Such interference is expected to be used to design useful functional devices such as various unitary transformers. It can also be used to design mode multiplexer/demultiplexers as will be shown in the latter chapters.

2.1 All Optical Discrete Fourier Transform Based on Multimode Interference Couplers

Part of the section reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 22(15), 1093–1095 (2010)] © IEEE [1].

Starting from this section, various unitary transformers based on the multi-mode interference effect will be discussed. For an example, an all optical discrete Fourier transform (DFT) device based on multimode interference (MMI) couplers is

proposed in this section. The DFT device is composed of a discrete sine transform (DST) device and a discrete cosine transform (DCT) device, both of which are based on MMI couplers. The main advantage of this DFT device is that it composes of two compact MMI couplers, which greatly reduces the fabrication effort on planar lightwave circuits (PLCs).

2.1.1 Introduction

Orthogonal frequency division multiplexing (OFDM), which implements discrete Fourier transform (DFT) as the coding algorithm, can significantly improve the system performance under the influence of chromatic dispersion and has been paid special attention by the research community [2–4]. As is mentioned by the researchers, the key coding process, i.e., DFT can be realized by the fast Fourier transform (FFT) algorithm. It is, however, still time consuming as it is done in the electronic domain. One may significantly save the time cost by the optical approach, such as the PLC based optical circuit [5], which consists of cascaded connected couplers and phase shifters. The structure of the PLC is relatively complex and is difficult to fabricate. The concept of the fast Fourier transform (FFT) can help to simplify the structure [6] and has been realized by Koichi Takiguchi et al. [7], but many couplers and phase shifters are still required.

A tentative approach to realize DFT is by the MMI couplers [8], which has already been used to realize functional devices, such as the 4*4 Hadamard transform [9, 10], the DST and the DCT [11]. The proper design of the MMI coupler based devices requires the knowledge of the transform matrix of the MMI coupler which relates the input port optical beams to the output ports optical beams [11, 12]. One may conclude from the expression of the transfer matrix that orthogonal transformers can be realized by properly placing the phase shifters at the input and output ports. In this way, we propose a new approach to realize DFT all optically.

2.1.2 Theory

2.1.2.1 Review of the DST/DCT IV Device Based on MMI Couplers [11]

The orthogonal transform DST and DCT could take different forms [13]. One of its kind can be realized directly by one single MMI coupler [11]. The results are essential to form the DFT and we revisit the derivation in the following. To start with, the input ports should be located at the following positions:

$$x = \left(i - \frac{1}{2}\right) \frac{W_e}{N} \quad (i = 1 \dots N) \quad (2.1)$$

where W_e is the effective width of the MMI coupler, N the total number of the input/output, i stands for the i th input/output port. It has been derived in [11] that the length of the MMI coupler should be properly set as $L_N = \frac{2nW_e^2}{N\lambda_0}$ where n is the refractive index of the waveguide, λ_0 the free space wavelength. In this way, the N input images of the beam will be converted as the N output images of the beam. These inputs and outputs are linearly related with each other. One may further add some phase shifters at the input and output ports, which will adjust the transfer matrix element [11, 12]. Under such conditions, the input vector and the output vector can be related by:

$$\mathbf{T} = \mathbf{D}_{out} \mathbf{V} \mathbf{B} \mathbf{V}^T \mathbf{D}_{in} \quad (2.2)$$

where $\mathbf{D}_{in}$ and $\mathbf{D}_{out}$ are the matrixes related to the phase shifter arrays at the input and output ports, $\mathbf{V}$ is an orthogonal matrix describing the relationship between the modes of the MMI coupler and the input and output beams. $\mathbf{V}$ has the following elements at the column i and row m [9]:

$$V_{in} = \begin{cases} \sqrt{\frac{2}{N}} \sin\left(\frac{n\pi}{N}(i-1/2)\right) & (n \neq N) \\ \sqrt{\frac{1}{N}} \sin(\pi(i-1/2)) & (n = N) \end{cases} \quad (2.3)$$

If the phase shifters have the phase shift of $\pi \frac{(i-1/2)^2}{2N}$ at the corresponding input and output ports, the elements of $\mathbf{T}$ will be adjusted as the elements of DST IV matrix [12]:

$$T_{lk} = j \exp\left(j\frac{\pi}{4}\right) \sqrt{\frac{2}{N}} \sin\left(\frac{\pi(l-1/2)(k-1/2)}{N}\right) \quad (2.4)$$

While DST can be realized all optically, DCT IV can be realized on the same MMI coupler by changing the labeling order and adding a phase shifter π at the even input ports.

2.1.2.2 Realization of the DFT by Combining the DST and DCT Devices

Neglecting the constant phase in the above equation, we have the MMI coupler based DCT IV and DST IV matrices to be with the elements of:

$$\begin{aligned} M_{DST}(l, k) &= \sqrt{\frac{2}{N}} \sin\left(\frac{\pi(l-1/2)(k-1/2)}{N}\right) \\ M_{DCT}(l, k) &= \sqrt{\frac{2}{N}} \cos\left(\frac{\pi(l-1/2)(k-1/2)}{N}\right) \end{aligned} \quad (2.5)$$

We may combine the two MMI couplers to form a DFT matrix by adding 2*2 directional couplers. In this case, we have:

$$\begin{aligned}
M(l, k) &= \sqrt{\frac{1}{N}} \exp\left(-j \frac{\pi(l-1/2)(k-1/2)}{N}\right) \\
&= \sqrt{\frac{1}{N}} \exp\left(-j \frac{\pi(l-1)(k-1)}{N}\right) \exp\left(-j \frac{\pi(l-1+k-1)}{2N}\right) \exp\left(-j \frac{\pi}{4N}\right)
\end{aligned} \tag{2.6}$$

The matrix can be further adjusted by adding more shifters at the input and output ports, and the transfer matrix will be modified as:

$$T = D_{out} M D_{in} \tag{2.7}$$

where D_{in} and D_{out} are the diagonal matrixes relating to the phase shifters, whose phase shift is $\frac{(i-1)\pi}{2N}$ at the i th port of the input and the output. In this case, the transfer matrix will have the element as:

$$T(l, k) = \exp\left(-j \frac{\pi}{4N}\right) \sqrt{\frac{1}{N}} \exp\left(-j \frac{\pi(l-1)(k-1)}{N}\right) \tag{2.8}$$

Considering the odd rows of the matrix, it coincides with the rows of the DFT matrix:

$$T(2l-1, k) = \exp\left(-j \frac{\pi}{4N}\right) \sqrt{\frac{1}{N}} \exp\left(-j \frac{2\pi(l-1)(k-1)}{N}\right) \tag{2.9}$$

This means that if $x(n)$ is injected into the input of the device, its DFT $X(k)$ can be obtained at the odd ports of the output, if $k = 0 \dots N/2-1$.

For the latter $N/2$ points with $k = N/2 \dots N-1$, one may exploit the property of DFT:

$$\begin{aligned}
k &= k' + \frac{N}{2} \\
X\left(k' + \frac{N}{2}\right) &= \sum_{n=0}^{N-1} (-1)^n x(n) \exp\left(-j \frac{2\pi n k'}{N}\right) \\
k' &= 0 \dots \frac{N}{2} - 1
\end{aligned} \tag{2.10}$$

The property suggests that the latter $N/2$ points $k = N/2 \dots N-1$ can be transformed to the former $N/2$ points $k' = 0 \dots N/2-1$ if the phase shift of π is introduced at the even input. Hence, if the transfer matrix $\mathbf{T}$ can be converted to transfer matrix $\mathbf{T}' = \mathbf{T}\mathbf{Q}$, where the matrix $\mathbf{Q}$ has the following form:

$$\mathbf{Q} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & 1 & \\ & & & -1 \\ & & & & \ddots \end{pmatrix}$$

One may realize the DFT. The key problem to introduce the matrix $\mathbf{Q}$. It can be done by adding more phase shifters, or one may adopt the concept proposed by Zhongde Wang [13], i.e., the transform matrixes for the DST IV and the DCT IV are related by the following equation:

$$\begin{aligned} \mathbf{M}_{\text{DCT}} &= \mathbf{P}\mathbf{M}_{\text{DST}}\mathbf{Q} \\ \mathbf{M}_{\text{DST}} &= \mathbf{P}\mathbf{M}_{\text{DCT}}\mathbf{Q} \end{aligned} \quad (2.11)$$

where matrix $\mathbf{P}$ has the following form:

$$\mathbf{P} = \begin{pmatrix} & & & 1 \\ & & & & 1 \\ & & \ddots & & & 1 \\ & & & 1 & & \\ & 1 & & & & \\ 1 & & & & & \end{pmatrix}$$

As is shown earlier, the $\mathbf{M} = \frac{1}{\sqrt{2}}(\mathbf{M}_{\text{DCT}} - j\mathbf{M}_{\text{DST}})$ is obtained by adding a few 2*2 couplers at the output. We may consider the other output port of the 2*2 coupler, which has the transfer matrix as:

$$\begin{aligned} \mathbf{M}' &= \frac{1}{\sqrt{2}}(\mathbf{M}_{\text{DST}} - j\mathbf{M}_{\text{DCT}}) \\ &= \frac{1}{\sqrt{2}}\mathbf{P}(\mathbf{M}_{\text{DCT}} - j\mathbf{M}_{\text{DST}})\mathbf{Q} \\ &= \mathbf{P}\mathbf{M}\mathbf{Q} \end{aligned} \quad (2.12)$$

If the input phase shifters $\frac{(i-1)\pi}{2N}$ to obtain the transfer matrix $\mathbf{T}$ is added, the other output ports of the 2*2 couplers will have the transfer matrix of $\frac{1}{\sqrt{2}}(\mathbf{M}_{\text{DST}} - j\mathbf{M}_{\text{DCT}})\mathbf{D}_{in}$. If the labeling order of other ports of the couplers are inverse with the respect to the ports to generate matrix $\mathbf{T}$ and the phase shifters $\frac{(i-1)\pi}{2N}$ are added in the inverse order as well, the transfer matrix $\mathbf{T}'$ will be:

$$\begin{aligned} \mathbf{T}' &= \mathbf{P}\mathbf{D}'_{out}\mathbf{P}\mathbf{M}\mathbf{Q}\mathbf{D}_{in} \\ &= \mathbf{D}_{out}\mathbf{M}\mathbf{D}_{in}\mathbf{Q} \\ &= \mathbf{T}\mathbf{Q} \end{aligned}$$

$$D'_{out}(i, i) = \left(\frac{N - i + 1}{2N} \right) \pi \quad (2.13)$$

In this way, we will get the latter half of the DFT $X(k)$ $k = N/2 \dots N-1$ at the other ports of the 2×2 couplers with the odd labeling number.

Therefore, we have proposed a DFT device by combining a DST device and a DCT device with N 2×2 couplers. It should be noted that phase shifters should be added to adapt the transfer matrix to the DFT matrix.

2.1.3 Simulation Results and Discussion

The proposed all optical DFT is tested on two 8-port MMI couplers, which are used to realize the DST device and the DCT device. Each MMI coupler has the index of 3.473 and the background index of 3.455 and its width is $72 \mu\text{m}$. The wavelength of the signal is 1064nm . It can be evaluated that the effective width is $73 \mu\text{m}$ and the length of the MMI coupler should be $4344 \mu\text{m}$. The 8 input and output ports of the MMI couplers are composed of 8 waveguides with the width of $2.6 \mu\text{m}$. These input and output waveguide located at the positions indicated by Eq. (2.1). The index and the wavelength parameters are from the published literatures [12, 14]. The fundamental modes of the 8 input waveguides are launched into the MMI couplers with different amplitudes and phases.

In order to realize DST/DCT device, phase shifters should be added as suggested in our previous work [11]. Furthermore, as discussed in the previous section, the phase shift $\frac{(i-1)\pi}{2N}$ should be added. In addition to that, 2×2 couplers are added to combine the DST and the DCT devices, with a group of ports having the output in the form of $\frac{1}{\sqrt{2}}(a-jb)$ and the other group of ports having the output in the form of $\frac{1}{\sqrt{2}}(b-ja)$. By adding the phase shift of $\frac{(i-1)\pi}{2N}$ at one group output, and the inverse ordered phase shifter $\frac{(i-1)\pi}{2N}$ at the other group output, one gets the DFT results of $X(0)$, $X(1)$, $X(2)$ and $X(3)$ at the odd output ports at one group output, and $X(4)$, $X(5)$, $X(6)$ and $X(7)$ at the other group output. The optical schematic is shown in Fig. 2.1.

The beam propagation inside the MMI coupler is simulated by the two dimensional beam propagation method (BPM).

The validity of the DFT device is tested with the input vector as $(11,111,111)^T$, $(11,110,000)^T$, $(11,000,000)^T$, and $(1,000,000)^T$. The resulting amplitudes of the output vectors are shown in Fig. 2.2. It can be clearly seen that the BPM simulation results match the ideal DFT results perfectly. The deviation between the two is less than 10%.

Afterwards, the fabrication error impact on the performance is investigated. The robustness of the device is tested by assuming 1% fabrication error in length ($40 \mu\text{m}$ length error) will result in 10% output amplitude deviation with the input vector as $(1,000,000)^T$. Increasing the robustness of the device is an important problem, which

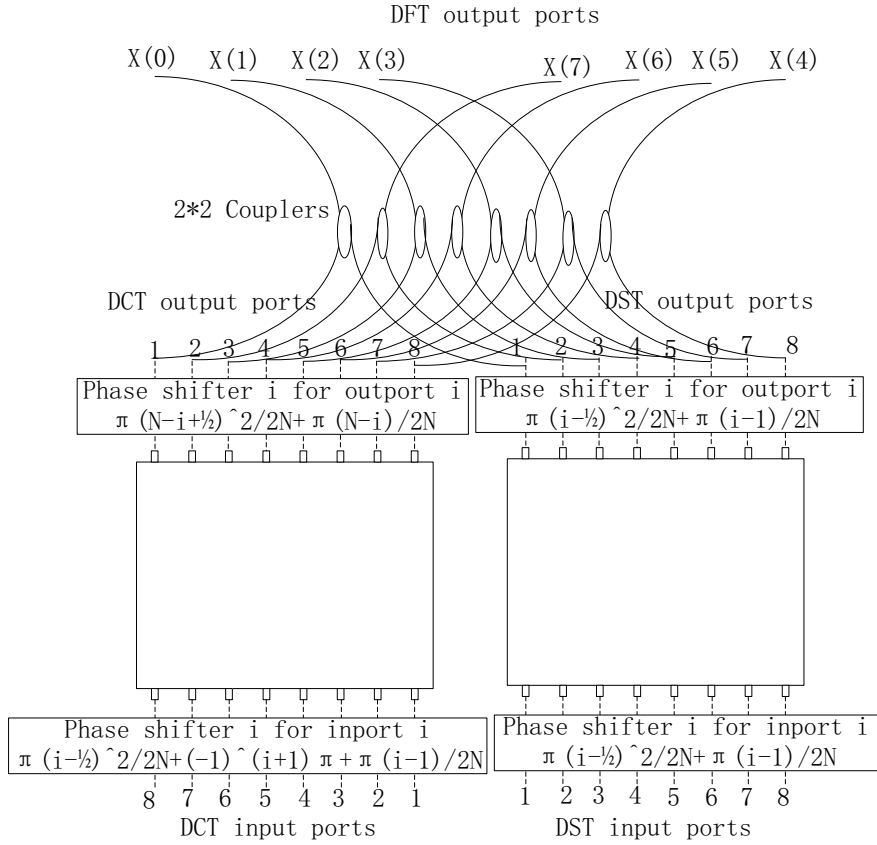
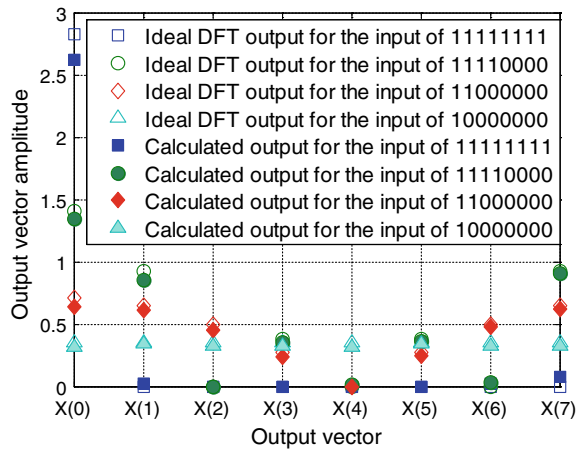


Fig. 2.1 Optical schematic of MMI coupler based DFT device. The figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 22(15), 1093–1095 (2010)] © IEEE [1]

Fig. 2.2 Output vectors obtained by the ideal DFT and BPM simulation. The figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 22(15), 1093–1095 (2010)] © IEEE [1]



can be achieved by adjusting the refractive index and the geometrical parameters of the MMI couplers.

2.1.4 Summary

A novel device to realize all optical DFT based on the MMI couplers combined by a few 2×2 directional couplers. Compared with other approaches, the proposed device is compact and easy to fabricate. It is expected that the proposed device will play an important role in all optical signal processing, particularly in the future optical OFDM systems.

2.2 Arbitrary Periodical Pattern Generation Based on Fractional Talbot Effect

In this section, the study on multimode interference is extended to free space. A novel approach to realize all-optical discrete Fourier transform (DFT) in free space is proposed for the first time to the best of our knowledge. The new method is based on the fractional Talbot effect, which alters the phases and the amplitudes of the periodically distributed beams without changing their common beam shape. A comprehensive theoretical framework for the DFT realization in free space is rigorously presented based on the Fresnel diffraction theory. Numerical simulations accompany the theory to provide a detailed analysis. Experimental demonstrations are conducted via a periodically modulated spatial light modulator (SLM) illuminated by the plane wave, and an excellent agreement is observed compared with the theoretical predictions.

2.2.1 Introduction

Discrete Fourier transform (DFT) is one of the most frequently used tools for signal processing, filtering and coding [15–17]. For decades, realization of DFT either numerically or physically has been an important research topic for the industrial and scientific communities.

A generally implemented way to realize DFT is based on the digital signal processing (DSP) technique. Cooley and Tukey [18] have proposed to split the sequence into two groups for the even terms and the odd terms during the transform. Such treatment divides the operation of an N -point DFT into two $N/2$ -point DFTs. Further division of the DFT operation brings up with the well-known fast Fourier transform (FFT) algorithm, which has been implemented in most of the

modern communication systems. While the FFT algorithm has been successful, it still requires significant computational resources. An N -point FFT requires about $N \log_2(N)$ complex multiplications, which is not trivial for the long sequences. In optical communication systems, the computational cost becomes even more prominent, particularly for the high-speed transmission systems, e.g., the 400 Gb/s optical orthogonal frequency division multiplexing (OFDM) systems [19].

Researchers have been searching for the alternative ways to realize DFT more efficiently, especially through the all-optical means. Marhic and Takiguchi have proposed to realize DFT by couplers and phase shifters [6, 7]. A more compact and efficient approach has been proposed by Kang by exploiting the multimode interference (MMI) effect inside optical waveguides [20, 21]. Apart from the spatial domain DFT studies [6, 7, 20, 21], the time domain DFT has been proposed to be realized based on the temporal MMI effect [22]. Up to date, there have been no discussions on the realization of DFT through free space optics. A vacancy for the related research is partially because the beam transmission in free space obeys the Fresnel diffraction law and the far field pattern is automatically the continuous Fourier transform of the input pattern. However, the DFT is substantially different from the continuous Fourier transform. An ideal DFT scheme should only alter the discrete phases and the amplitudes of the periodically distributed beams without changing their common beam shape, which is not possible for the continuous Fourier transform. Such ideal DFT scheme can be quite useful for beam shaping, light manipulation and all optical signal processing, particularly in the near field regime.

For the first time to the best of our knowledge, it is revealed in this study that Talbot effect, which has been studied intensively for almost 180 years [23–28], is potentially capable of realizing all-optical DFT if the propagation distance is set as $1/N$ of the standard Talbot distance. While it is well accepted that the fractional Talbot effect occurs under such set-up, it has not been reported that DFT is possible to be realized in this way. In this section, we explore the detailed theoretical framework from the traditional Fresnel diffraction theory and provide a rigorous proof that DFT does exist among the input and the output patterns provided that a quadratic phase term is added. The detailed simulation results show an excellent agreement with the theoretical predictions. Experiments are conducted by measuring the diffraction pattern from a spatial light modulator (SLM) with a periodical phase pattern to verify the newly proposed DFT realization approach, and the measurements agree well with the theoretical predictions.

2.2.2 Theory

The theoretical framework for all-optical DFT in free space includes the one-dimensional DFT, the two-dimensional DFT for the periodical Dirac delta pulses and the extension of the above concepts for the periodical beam patterns with an arbitrary common beam shape. The derived results, particularly the one-dimensional DFT in free space, resembles the results of temporal MMI effect based DFT in [22].

However, the theoretical framework in [22] is not possible to be extended to the two-dimensional DFT case.

2.2.2.1 The Talbot Effect Based One-Dimensional DFT Scheme for the Dirac Delta Pulses

The scalar wave propagation in free space is governed by the scalar Holmholtz equation:

$$\frac{\partial^2 \psi}{\partial z^2} + \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + k^2 \psi = 0 \quad (2.14)$$

where ψ is the scalar wave function. Assuming the beam propagates in the z direction, and it remains constant in the y direction, we have the following

$$\begin{aligned} \psi &= \phi \exp(jkz) \\ \frac{\partial^2 \phi}{\partial z^2} + 2jk \frac{\partial \phi}{\partial z} + \frac{\partial^2 \phi}{\partial x^2} &= 0 \end{aligned} \quad (2.15)$$

where ϕ is the slowly varying envelope. Under Fresnel diffraction theory or the paraxial approximation, one has

$$2jk \frac{\partial \phi}{\partial z} + \frac{\partial^2 \phi}{\partial x^2} = 0 \quad (2.16)$$

Assuming that there are periodical Dirac delta pulses on the input plane, one has

$$\begin{aligned} \phi(0, x) &= \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{N-1} a_{nN+k} \delta(x - (nN + k)\Lambda) \\ a_{nN+k} &= a_k \end{aligned} \quad (2.17)$$

where δ is the Dirac delta function, Λ the spatial separation between the adjacent pulses, N the period of the pulse chain, $N\Lambda$ the total spatial period of the pattern, a_k the periodical coefficient of the pulse chain. According to the property of the Dirac delta function, one has

$$\begin{aligned}
\phi(0, x) &= \frac{1}{N\Lambda} \sum_{n=-\infty}^{+\infty} A_n \exp(jn\kappa x) \\
A_n &= \sum_{k=0}^{N-1} a_k \exp\left(-jnk \frac{2\pi}{N}\right) \\
\kappa &= \frac{2\pi}{N\Lambda}
\end{aligned} \tag{2.18}$$

where A_n is a periodical coefficient with the period as N , κ the fundamental spatial frequency. After propagation for the distance of z , one has [26]

$$\phi(z, x) = \frac{1}{N\Lambda} \sum_{n=-\infty}^{+\infty} A_n \exp(jn\kappa x) \exp(j\beta_n z) \tag{2.19}$$

where β_n is the propagation constant for the spatial frequency $n\kappa$. Substituting Eq. (2.19) into (2.16), one gets,

$$\beta_n = -\frac{1}{2k} (n\kappa)^2 = -\lambda\pi \left(\frac{n}{N\Lambda}\right)^2 \tag{2.20}$$

Since the beam has the spatial period of $N\Lambda$ in the x direction, naturally, we have the Talbot distance to be [26]

$$z_{Talbot} = \frac{N^2 \Lambda^2}{\lambda} \tag{2.21}$$

where the standard self-imaging effect takes place.

Assuming the propagation distance is $z_T = z_{Talbot}/N$, where the fractional Talbot effect occurs, we have the beam to evolve as

$$\begin{aligned}
z_T &= \frac{z_{Talbot}}{N} = \frac{N\Lambda^2}{\lambda} \\
\phi(z_T, x) &= \frac{1}{N\Lambda} \sum_{n=-\infty}^{+\infty} A_n \exp(jn\kappa x) \exp\left(-j\lambda\pi \left(\frac{n}{N\Lambda}\right)^2 z_T\right) \\
&= \frac{1}{N\Lambda} \sum_{n=-\infty}^{+\infty} A_n \exp(jn\kappa x) \exp\left(-j\frac{n^2\pi}{N}\right)
\end{aligned} \tag{2.22}$$

Since A_n is periodical, one has

$$A_{IN+m} = A_m \tag{2.23}$$

Using (2.23) and (2.18), Eq. (2.22) may be converted to

$$\begin{aligned}
\phi(z_T, x) &= \sum_{l=-\infty}^{+\infty} \delta(x - l\Lambda) \left(\frac{1}{N} \sum_{m=0}^{N-1} A_m \exp\left(-j \frac{m^2 \pi}{N}\right) \exp\left(\frac{j 2\pi m l}{N}\right) \right) \\
&= \sum_{l=-\infty}^{+\infty} C_l \delta(x - l\Lambda) \\
C_l &= \frac{e^{-j \frac{\pi}{4}}}{\sqrt{N}} \exp\left(j \frac{l^2 \pi}{N}\right) \sum_{k=0}^{N-1} a_k \exp\left(j \frac{k^2 \pi}{N}\right) \exp\left(-j \frac{2lk \pi}{N}\right) \quad (2.24)
\end{aligned}$$

Obviously, C_l and a_k have a relationship of DFT despite a constant phase term and two quadratic phase terms $\exp\left(j \frac{l^2 \pi}{N}\right)$ and $\exp\left(j \frac{k^2 \pi}{N}\right)$. Therefore, the N periodical pulses on the input plane become N periodical pulses at $z = z_T$, with the coefficients for the input and the output pulses being the DFT pairs.

2.2.2.2 The Talbot Effect Based Two-Dimensional DFT Scheme for the Dirac Delta Pulses

It is easy to generalize the above derivation for the two-dimensional case to realize the two-dimensional DFT. In this case, Eq. (2.16) will be

$$2jk \frac{\partial \phi}{\partial z} + \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \quad (2.25)$$

We may assume the input two-dimensional Dirac delta pulse chain has the equal period $N\Lambda$ in the x direction and the y direction:

$$\begin{aligned}
\phi(0, x, y) &= \sum_{m=-\infty}^{+\infty} \sum_{l=0}^{N-1} \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{N-1} a_{nN+k, mN+l} \\
&\quad \delta(x - (nN + k)\Lambda) \delta(y - (mN + l)\Lambda) \\
a_{nN+k, mN+l} &= a_{k,l} \quad (2.26)
\end{aligned}$$

The field at $z_T = z_{\text{Talbot}}/N$ will be

$$\begin{aligned}
\phi(z_T, x, y) &= \sum_{p=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} C_{p,q} \delta(x - p\Lambda) \delta(y - q\Lambda) \\
C_{p,q} &= \frac{-j}{N} \exp\left(j \frac{(p^2 + q^2)\pi}{N}\right) \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} a_{k,l} \\
&\quad \exp\left(j \frac{(k^2 + l^2)\pi}{N}\right) \exp\left(-j \frac{2(lp + kq)\pi}{N}\right) \quad (2.27)
\end{aligned}$$

which suggests $C_{p,q}$ and $a_{k,l}$ are two-dimensional DFT pairs despite the quadratic phase terms.

2.2.2.3 Extension to the Arbitrary Beam Forms

Here we use the two-dimensional case as an example and it is easy to be simplified to the one-dimensional case. To extend the above results for the periodical Dirac delta pulses to the case of the arbitrary beam forms, one has to assume that the periodical beam on the input plane has a discrete amplitude/phase distribution pattern, i.e., the beam shape remains identical within the range of $\Lambda \times \Lambda$ in the x direction and the y direction while each beam possesses a distinctive amplitude/phase. The total spatial period is $N\Lambda \times N\Lambda$. In this case, one has

$$\begin{aligned} \phi(0, x, y) = & g(x, y) \otimes \sum_{m=-\infty}^{+\infty} \sum_{l=0}^{N-1} \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{N-1} a_{nN+k, mN+l} \\ & \delta(x - (nN + k)\Lambda) \delta(x - (mN + l)\Lambda) \end{aligned} \quad (2.28)$$

where $\otimes$ denotes the convolution operation, $g(x, y)$ the common beam shape. The system under the Fresnel diffraction theory is a linear system, and therefore, one has the pattern on the output plane to be

$$\begin{aligned} \phi(z_T, x, y) &= \left(g(x, y) \otimes \sum_{m=-\infty}^{+\infty} \sum_{l=0}^{N-1} \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{N-1} a_{nN+k, mN+l} \delta(x - (nN + k)\Lambda) \delta(x - (mN + l)\Lambda) \right) \\ &\otimes G(z_T, x, y) \end{aligned} \quad (2.29)$$

where $G(z_T, x, y)$ is the Green's function with respect to the distance of z_T . According to the property of the convolution operation, we have

$$\begin{aligned} \phi(z_T, x) &= g(x, y) \otimes \left(\sum_{m=-\infty}^{+\infty} \sum_{l=0}^{N-1} \sum_{n=-\infty}^{+\infty} \sum_{k=0}^{N-1} a_{nN+k, mN+l} \right. \\ &\quad \left. \delta(x - (nN + k)\Lambda) \delta(x - (mN + l)\Lambda) \otimes G(z_T, x, y) \right) \\ &= g(x, y) \otimes \left(\sum_{p=-\infty}^{+\infty} \sum_{q=-\infty}^{+\infty} C_{p,q} \delta(x - p\Lambda) \delta(x - q\Lambda) \right) \end{aligned} \quad (2.30)$$

Clearly, the above expression represents a periodical waveform (possessing a common beam shape within the range of $\Lambda \times \Lambda$) with a discrete amplitude/phase distribution pattern on the output plane at the distance of z_T . Therefore, the DFT

scheme presented in this section can operate for the beams with an arbitrary common beam shape.

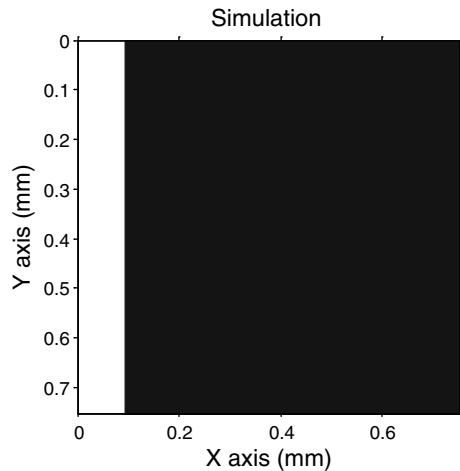
2.2.3 Numerical Results with Infinite Periodical Patterns

The proposed free space all-optical DFT realization scheme is tested by the numerical simulations. In the simulation, the periodical pattern is assumed to be infinite so that the ideal DFT can be realized. Such infinite pattern is easily realizable in the simulation via the Fourier transformation of one spatial period. The beam shape within the range of Λ is assumed to be a constant value, and the amplitudes of the periodical beams remain 1. The phase within the range of Λ can be adjusted accordingly. Such set-up is easy to be realized experimentally via a phase tunable SLM.

The pixel for phase tuning is $9.4 \mu\text{m} \times 9.4 \mu\text{m}$. In the x direction, 10 pixels form a beam with a constant phase and 8 beams form a spatial period, i.e., $N\Lambda = 8 \times 10 \times 9.4 \mu\text{m} = 0.75 \text{ mm}$. In the y direction, the phase remains constant. According to the analysis in Sect. 2.5.2, such arrangement will create a 8-point DFT pattern at $z = z_T$. In the simulation, the input beams are loaded with the quadratic phase term $\exp(jk^2\pi/N)$. Therefore, the input vector should be $[1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$, and hence, the output vector should be the DFT of it, i.e. $[2\sqrt{2} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$ the output pattern should have the first beam intensity to be 8, while the rest 7 beam-intensities are 0. The intensity for one period of the beam at $z = z_T$ is plotted in Fig. 2.3, which shows the results as expected.

Figure 2.4 is plotted to demonstrate how the beam evolves in the z direction. It can be seen from the figure that the intensities are identical for the 8 beams at $z = 0$. However, the intensity concentrates gradually towards the 1st beam, and the

Fig. 2.3 The Talbot effect based 1D DFT output. The input vector is $[1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$ and the output vector is $[2\sqrt{2} \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$. Intensity within one period is demonstrated here



concentration reaches its maximum at $z = z_T$. The figure clearly demonstrates how the all-optical DFT is realized via the fractional Talbot effect.

The two-dimensional DFT is also demonstrated numerically. The configuration is similar with respect to the one-dimensional DFT case. In the x direction, 10 pixels form a beam with a constant phase and 8 beams form a period. In the y direction, the same configuration is adopted. Therefore, a two-dimensional period of $0.75 \text{ mm} \times 0.75 \text{ mm}$ is contained within the pattern. Similarly, only phase changes for each beam. We assume the input beam is applied with the two-dimensional quadratic phase term, and the input vector should be $[11111111]^T \otimes [11111111]^T$, where $\otimes$ denotes the matrix Kronecker product. The two-dimensional DFT result should be $\left[2\sqrt{20000000}\right] \otimes \left[2\sqrt{20000000}\right]^T$. The results are plotted in Fig. 2.5 and a perfect match with the theory is observed.

Fig. 2.4 Evolution of the 1D Talbot effect based DFT scheme along the z axis. Intensity within one period is demonstrated here

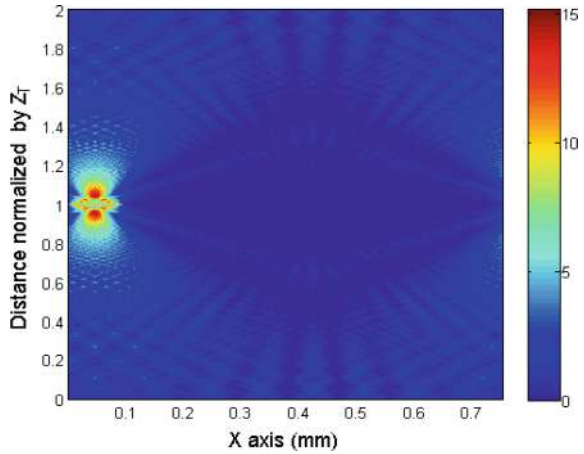
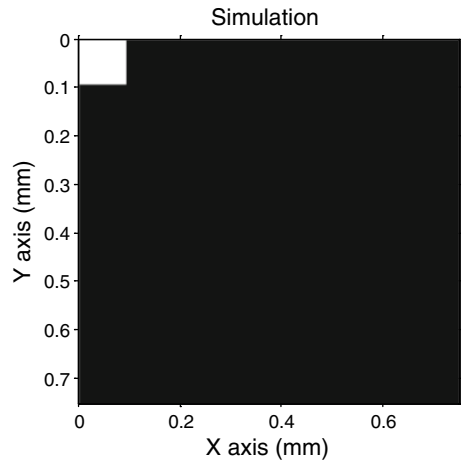


Fig. 2.5 The Talbot effect based 2D DFT output. The input vector is $[1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1] \otimes [1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]^T$ and the output vector is $\left[2\sqrt{20000000}\right] \otimes \left[2\sqrt{20000000}\right]^T$. Intensity within one period is demonstrated here



2.2.4 Experimental Verifications

The proposed free space all-optical DFT realization scheme is verified experimentally. A continuous wave laser at the wavelength of 671 nm is used as the light source. The laser beam is magnified by a laser-beam-expander. After magnification, the beam has the radius of about 10 mm. The laser beam is used to illuminate a SLM with the pixel size of $9.4 \mu\text{m} \times 9.4 \mu\text{m}$, and the SLM has the grid as 800×600 . The laser beam has its beam waist fully covering the SLM active area. Since the periodical pattern on the SLM is not infinite, irregularity might arise, particularly at the edge of the SLM. The diffracted beam is measured by a Complementary Metal Oxide Semiconductor (CMOS) camera at the distance of z_T from the SLM. The experimental setup is shown in Fig. 2.6.

First of all, the results in Sect. 2.5.3 are generated experimentally. Since the periodical pattern is not infinite on the SLM, the simulation is re-conducted to adapt to the real parameters. In the x direction of the SLM with 800 pixels, 10 pixels of the SLM form a beam with a constant phase and 8 beams form a period, i.e., $N\Lambda = 8 \times 10 \times 9.4 \mu\text{m} = 0.75 \text{ mm}$. Totally, 10 periods are formed in the x direction of the SLM. z_T is calculated as 105.3 mm and is slightly adjusted during the experiment. For the one-dimensional DFT, the phase in the y direction remains constant. The quadratic phase term is applied in the x direction. The simulated and the experimental intensity results at $z = z_T$ are shown in Fig. 2.7. For the two-dimensional DFT, the period of 0.75 mm is also applied in the y direction, which forms a 8×8 two-dimensional DFT scheme. The simulated and the experimental intensity results at $z = z_T$ for the two-dimensional DFT are shown in Fig. 2.8. Both figures demonstrate an excellent match between the theory and the experiment.

In order to further verify the correctness of the proposed all-optical DFT scheme, a different phase pattern is applied. While the input vector's amplitude remains 1, the phase is set as $[0 \ 0.5000 \ 0.3918 \ 0.8918 \ 0 \ 0.5000 \ 0.3918 \ 0.8918] \times \pi$, the

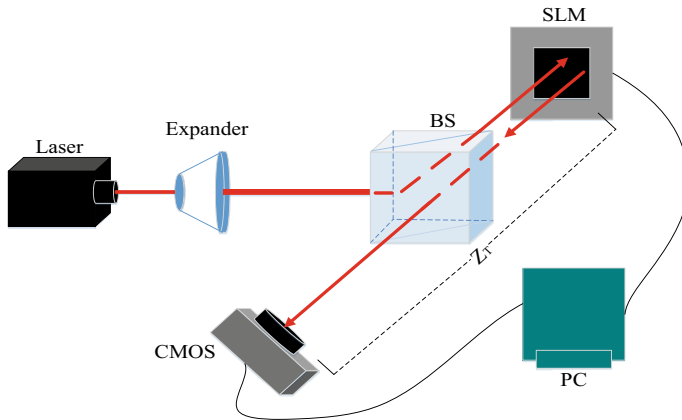


Fig. 2.6 The experimental set-up. BS: beam splitter, PC: personal computer

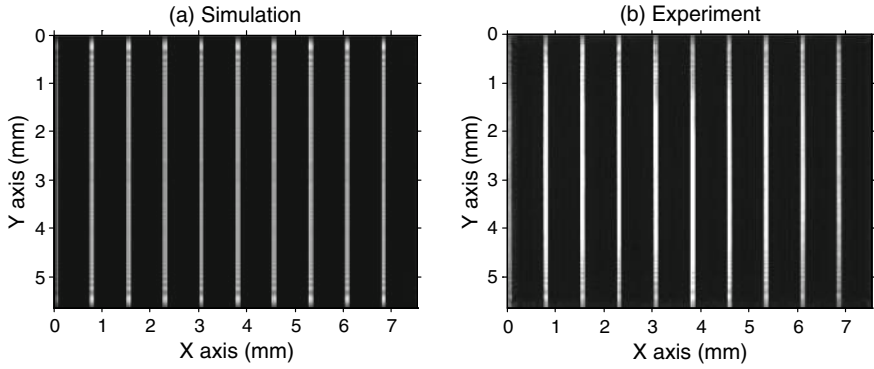


Fig. 2.7 Simulated (a) and experimental (b) results (intensity distribution) of the Talbot effect based 1D DFT output. The input vector is $[1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1]$ and the output vector is $[2\sqrt{2}0000000]$

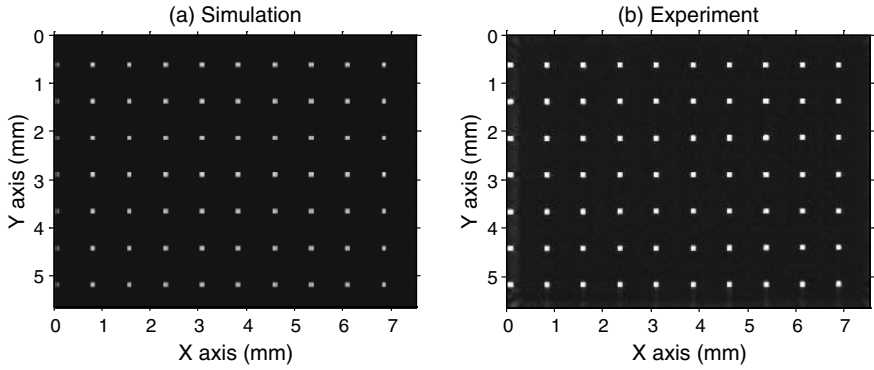


Fig. 2.8 Simulated (a) and experimental (b) results (intensity distribution) of the Talbot effect based 2D DFT output. The input vector is $[11111111] \otimes [11111111]^T$, and the output vector is $[2\sqrt{2}0000000] \otimes [2\sqrt{2}0000000]^T$

DFT of the above vector will be $[1.633\exp(j0.4459\pi) \ 0 \ 1.633\exp(-j0.3041\pi) \ 0 \ 1.633\exp(-j0.0541\pi) \ 0 \ 0 \ 0]$. Clearly, the DFT results have three non-zero terms and five zero terms. Such specific input phase distribution is computed from the well-known Gerchberg-Saxton algorithm [29]. Under such input phase pattern, one may expect three lines in each period within the intensity pattern for the one-dimensional DFT. The simulated and the experimental intensity results are shown in Fig. 2.9, which demonstrates the three lines within each period as expected. In order to study the intensity distribution within one period in a more accurate sense, the lateral (x directional) intensity distribution (by averaging 100 pixels in the middle of y direction) for the two periods in the middle of Fig. 2.9 are plotted in the one-dimensional sense in Fig. 2.10. In the figure, it can be concluded that the DFT has effectively taken

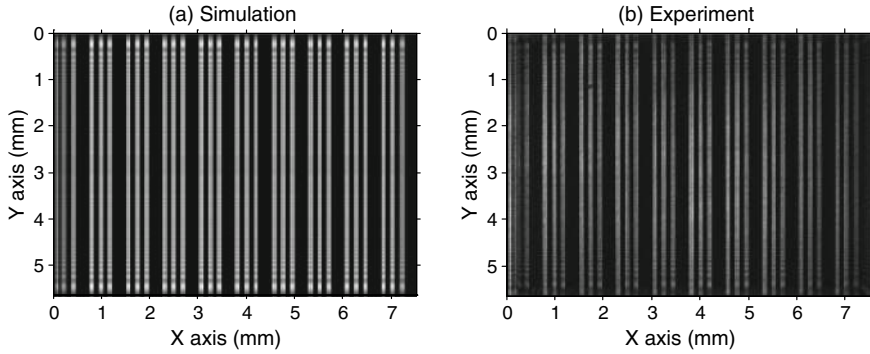


Fig. 2.9 Simulated (a) and experimental (b) results (intensity distribution) of the Talbot effect based 1D DFT output. The input vector's amplitude remains 1, the phase is set as $[0 \ 0.5000 \ 0.3918 \ 0.8918 \ 0 \ 0.5000 \ 0.3918 \ 0.8918] \times \pi$, the output vector is $[1.633 \exp(j0.4459\pi) \ 0 \ 1.633 \exp(-j0.3041\pi) \ 0 \ 1.633 \exp(-j0.0541\pi) \ 0 \ 0 \ 0]$

place and the three non-zero terms are in their exact locations, while the five zero terms are also correctly located. Although the intensities of the three non-zero terms are not exactly equal in the experiment, the deviation is within an acceptable range and it may be attributed to the experimental imperfections, such as the non-ideal plane wave during the illumination of the SLM.

The above specific one-dimensional pattern can also be extended to the two-dimensional case. The input phase pattern within the 8×8 period is the Kronecker product of the phase vector $[0 \ 0.5000 \ 0.3918 \ 0.8918 \ 0 \ 0.5000 \ 0.3918 \ 0.8918] \times \pi$ and its transpose. Therefore, the output pattern should have nine non-zero spots within each two-dimensional period of $0.75\text{mm} \times 0.75\text{mm}$, which is shown in Fig. 2.11. Again, the simulated results and the experimental measurements match perfectly.

2.2.5 Summary

In summary, a novel approach to realize all optical DFT in free space is proposed. The method is based on the fractional Talbot effect and a rigorous theoretical framework is established based on the Fresnel diffraction theory. Detailed numerical simulations are provided and an excellent agreement with the theory has been observed. Further validation of the proposed method is performed experimentally, through a SLM illuminated by the plane wave. The proposed method would be highly beneficiary for all-optical signal processing as well as other applications in free space.

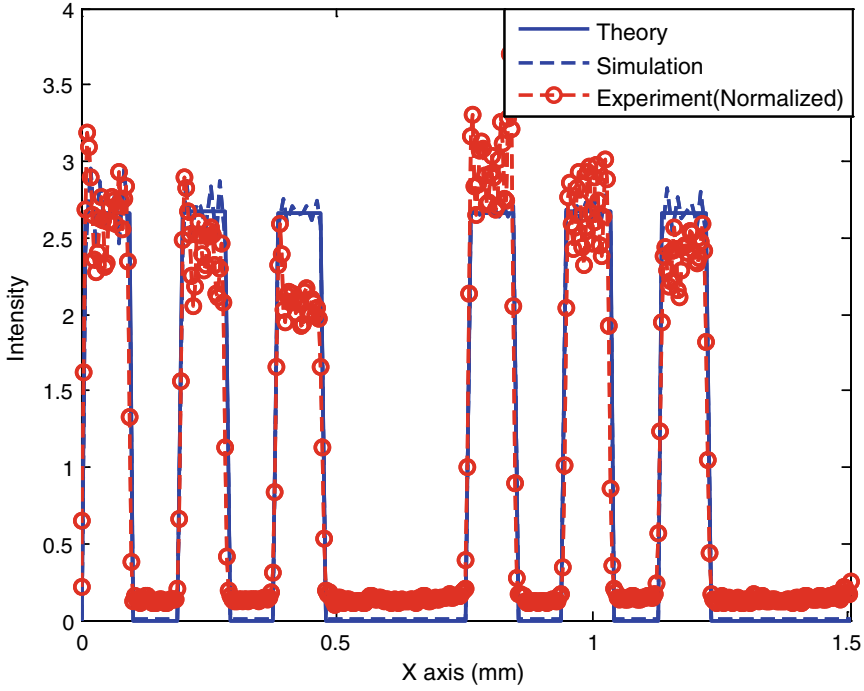


Fig. 2.10 The theoretical, simulated and experimental results (intensity distribution) of the Talbot effect based 1D DFT output plotted in the 1D sense for the two periods in the middle of x direction in Fig. 2.9. The input vector's amplitude remains 1, the phase is set as $[[0.5000 \ 0.3918 \ 0.8918 \ 0.5000 \ 0.3918 \ 0.8918] \times \pi]$, the output vector is $[1.633 \exp(j0.4459\pi) \ 0 \ 1.633 \exp(-j0.3041\pi) \ 0 \ 1.633 \exp(-j0.0541\pi) \ 000]$

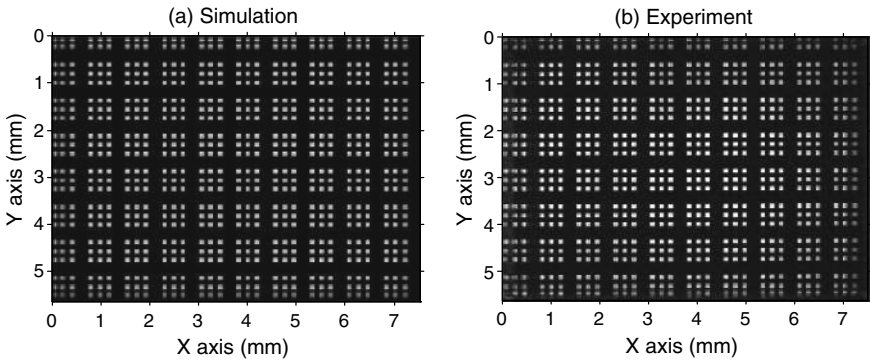


Fig. 2.11 Simulated (a) and experimental (b) results (intensity distribution) of the Talbot effect based 2D DFT output. The input vector's amplitude remains 1, and the phase within one period is the Kronecker product of the phase in Fig. 2.8

2.3 Ring Shape Talbot Effect

Part of the section reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30].

In this section, we report a novel type of Talbot effect in free space. While the traditional Talbot effect takes place in the lateral direction, the novel ring shape Talbot effect takes place for the beams with certain topological charges. A rigorous theoretical derivation for the new phenomenon is provided based on the Fresnel diffraction theory. Experimental demonstration is conducted by periodically modulating a spatial light modulator (SLM) to create the beams with the ring shape periodical phase distribution. With a properly adjusted phase pattern, the beam will exhibit equally spaced ring intensity distribution at the fractional Talbot length.

2.3.1 Introduction

Talbot effect, which is named after Talbot [23, 24], is a beam pattern re-emerging effect due to the periodical distribution of the index. It is caused by the periodical phase coincidence of the modes in waveguides or free space [28, 31–40].

In the optical waveguides, Talbot effect is also referred to as the multimode interference (MMI) effect [31–34], and the device is called the MMI couplers, which have been demonstrated in the first section of the chapter. The periodical distribution of the modes is assumed in the waveguides as the mode distribution fulfills the sinusoidal functions. In free space, periodical distributed phase mask will result in the periodical pattern as is shown in the previous section of the chapter. In both cases, Talbot effect exists in the lateral direction [24]. There could exist other forms of Talbot effects. For examples, Azaña et al. [25] suggested that Talbot effect could occur in the angular spectral domain if one use a point source which creates a spherical wave front with has a periodical field distribution in the angular domain. A spatial angular Talbot effect was reported by Samadian et al. [26] by converting the planar waveguide into the fan shape. The azimuth direction Talbot effect [27] can be realized by the interference among the Bessel beams. The azimuthally periodical grating suggested by Rasouli et al. [41] could also cause Talbot effect. The variants of Talbot effect have offered a new toolbox to design novel functional devices. For example, the spatial angular Talbot effect can be used to minimize the waveguide couplers [26].

As discussed above, Talbot effect could exist either in the lateral direction or in the angular direction [23–28, 31–41]. However, it has not been observed in the radial direction. If it is possible to have such a Talbot effect, one may create the equally spaced ring pattern, which might be of great importance to create radial gratings or other applications which requires ring-shape beam generation and manipulation. A straight forward example is the Bessel beam generation. It is well known that the Bessel beam can be generated by the ring shape bema in the spectral domain [42]. After diffractive propagation, the ring shape beam will turn into the non-diffractive

Bessel beam and could be potentially applied in the areas like bio-photonics [43, 44], free space communication [45], and ablation of metal thin films [46]. Atom trapping [47] and laser welding are other examples requires the creation and manipulation of the ring shape beams.

For the first time to the best of our knowledge, it is revealed in this section that the existence problem of the ring shape Talbot effect is solved. The condition to have such a novel phenomenon is to have a beam with a periodical distribution in the radial direction and a specific topological charge. A radially periodical phase mask with an angular vortex can be added to on a Gaussian beam to realize the above-mentioned conditions. We have established a comprehensive theoretical framework based on the paraxial beam propagation equation with the experimental measurements confirming the theoretical predictions.

2.3.2 Theory

The new phenomenon occurs in the periodical cylindrical coordinates, and the equation to describe it should be the Helmholtz equation in the cylindrical coordinates. To simplify the analysis, scalar wave is assumed:

$$\frac{\partial^2 \psi}{\partial z^2} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \varphi^2} + k^2 \psi = 0 \quad (2.31)$$

where ψ is the scalar wave function, r and φ the variables in the radial direction and the azimuth direction. It is usually assumed that slowly varying envelope ϕ is multiplied with a fast-oscillating phase term in the z direction

$$\psi(r, \varphi, z) = \phi(r, \varphi, z) \exp(jkz) \quad (2.32)$$

where k is the free space wave number. Since the propagation distance is much longer than the wavelength and the angular distribution is relatively small in size compared with the scale of the propagation distance, the paraxial approximation can be adopted, and one may have:

$$2jk \frac{\partial \phi}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \phi}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \phi}{\partial \varphi^2} = 0 \quad (2.33)$$

The equation can be further simplified by assuming that the beam can be decomposed into the radial distribution and the angular phase distribution, which is the property of the vortex beam. If it is the case, the slowly varying envelope can be written as $\phi = f(r)e^{jl\varphi}$, where l is the topological charge, and f a radial function. We have the equation for the radial distribution function f as:

$$2jk \frac{\partial f}{\partial z} + \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) - \frac{l^2}{r^2} f = 0 \quad (2.34)$$

Furthermore, f should be assumed as periodical to realize the radial Talbot effect and hence, we have:

$$f(0, r) = \sum_{n=-\infty}^{+\infty} \delta(r - n\Lambda) = \frac{1}{\Lambda} \sum_{n=-\infty}^{+\infty} \exp(jn\kappa r) \quad (2.35)$$

$$\kappa = \frac{2\pi}{\Lambda}$$

where δ is the Dirac delta function, Λ is the spatial period of the pattern in the radial direction, and κ is the fundamental spatial frequency in the radial direction. The spatial frequency $n\kappa$ corresponds to a specific spatial mode with a specific propagation constant β_n [24, 26].

$$f(z, r) = \frac{1}{N\Lambda} \sum_{n=-\infty}^{+\infty} \exp(jn\kappa r) \exp(j\beta_n z) \quad (2.36)$$

The propagation constant β_n can be computed from the equation for f . To simplify the computation process so that an analytical expression can be obtained, the following assumptions should be taken:

$$r \gg \Lambda$$

$$\beta_n^2 \gg \beta_n \frac{1}{r} \gg \frac{1}{r^2} \quad (2.37)$$

Under these assumptions, the equation for f can be simplified as:

$$2jk \frac{\partial f}{\partial z} + \frac{\partial^2 f}{\partial r^2} = 0 \quad (2.38)$$

which can be used to evaluate the propagation constant as

$$\beta_n = -\frac{1}{2k} (n\kappa)^2 = -\lambda\pi \left(\frac{n}{\Lambda} \right)^2 \quad (2.39)$$

where λ is the free space wavelength. Equation (2.39) suggests that the ring shape spatial modes are the same the lateral spatial modes, which can be gathered for the generation of the Talbot effect [24].

The condition to have the Talbot effect for such field distribution is to have the length $z_{\text{Talbot}} = \Lambda^2/\lambda$, where the phase matching condition meets [24] and the field can potentially re-emerge, i.e.

$$\beta_n z = -\lambda \pi \left(\frac{n}{\Lambda} \right)^2 \Lambda^2 / \lambda = -n^2 \pi \quad (2.40)$$

Hence, the self-imaging effect takes place, which creates the input image with the shift of half the period in the radial direction [24]. When $z = 2m z_{\text{Talbot}}$, where m is an integer number, the same image of the beam will reappear on the object plane, which is exactly the same as the input image albeit with the attenuation due to the diffraction effect.

The fractional Talbot effect will take place in case the propagation distance is $1/N$ of the fundamental Talbot length:

$$z_{TN} = \frac{z_{\text{Talbot}}}{N} = \frac{N \Lambda^2}{\lambda} \quad (2.41)$$

where N is an integer. Under the condition to have the fractional radial Talbot effect, the wave function becomes:

$$\phi(z_{TN}, r) = \frac{e^{-j\frac{\pi}{4}}}{\sqrt{N}} \sum_{p=-\infty}^{+\infty} \exp\left(j\frac{p^2\pi}{N}\right) \delta\left(r - p\frac{\Lambda}{N}\right) \quad (2.42)$$

It indicates that the fractional radial Talbot effect will create N times number of rings compared with the input pattern. The rings are with equal amplitudes and different phases $\exp(jp^2\pi/N)$, and they are equally spaced with the spatial separation as Λ/N . Inversely, we may expect that the same rings with the phase term of $\exp(-jp^2\pi/N)$ will converge to a single ring at the N -fractional Talbot length z_{TN} . This is due to the inverse propagation principle by taking the complex conjugate on the wave propagation equation.

2.3.3 Results and Discussions

The proposed radial Talbot effect has been proved theoretically and will be tested experimentally in this subsection. The experimental setup is as follows and is shown in Fig. 2.12. The light source is a 671 nm continuous wave laser. In order to achieve a near plane wave beam, the output of the light source is magnified by a laser-beam-expander so that the beam radius is about 10 mm, which is then projected to a SLM. The SLM has the grid of 800×600 with each pixel size as $9.4 \mu\text{m} \times 9.4 \mu\text{m}$. The effect area of the SLM is smaller than the beam area so that the beam resembles the plane wave. The SLM phase pattern is set as the ring shape pattern. The ring width is set as $94 \mu\text{m}$. The periodical distribution of the phase in the radial direction has the period as 4 rings and therefore $\Lambda = 9.4 \mu\text{m} \times 10 \times 4 = 0.376 \text{ mm}$. Each ring within the period has been applied with the quadratic phase term $\exp(-jp^2\pi/N)$. Such configuration ensures that the beam will merge from 4 rings to 1 ring within the period after propagating the fractional Talbot distance. Here, we have chosen the

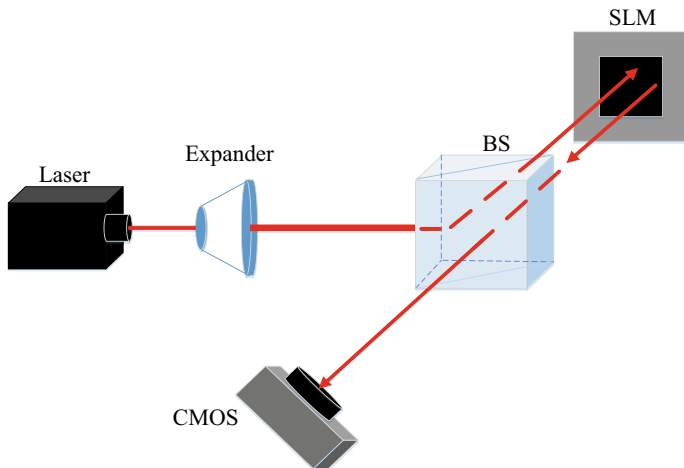


Fig. 2.12 The experimental set-up. BS: beam splitter. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

period number N to 4 and the fractional Talbot distance z_{TN} can be computed by the formula as 52.7 mm. A Complementary Metal Oxide Semiconductor (CMOS) camera located at multiple of the distance z_{TN} from the SLM is used to capture the output intensity distributions. The camera has the pixel size of $3.45 \mu\text{m} \times 3.45 \mu\text{m}$, and the total grid is 2160×4096 .

Firstly, we assume that the fundamental vortex mode with the topological charge of 0 is used. Since $N = 4$, the phase on the rings within the period can be calculated as $0, -\pi/4, -\pi, -9\pi/4$. Since the SLM can only be applied with the position phase pattern, these phases are readjusted as $0, 7\pi/4, \pi, 7\pi/4$. The corresponding SLM phase pattern is demonstrated in Fig. 2.13.

The CMOS camera measurement results are shown in Fig. 2.14b. To check the validity of the theory, the simulation results based on the paraxial propagation model is shown in Fig. 2.14a to compare with the experimental measurement results. In the figure, the central spot has been removed. Otherwise, it will be too bright so that the rest part of the figure will be shown in a very weak sense. The comparison between the two results demonstrates the validity of the theory. Although the patterns are identical, the amplitude or the intensity of the results might differ, because the CMOS camera will saturate if intensity reaches the threshold and creates a relatively flat distribution with the change of the radius while the simulation results suggest the intensity will decrease as the radius increases. The discrepancy in terms of intensity is well amended by the perfect match between the numerical and experimental results in terms of the positions of the rings. The distance between the rings in the measurement is $\Lambda = 0.376 \text{ mm}$, which is in exact agreement with the period of Λ . And this suggests that the fractional ring shape Talbot effect did occur.

During the analysis of the Talbot effect, the evolution of the beam along the propagation distance is usually demonstrated to show the self-imaging effect and

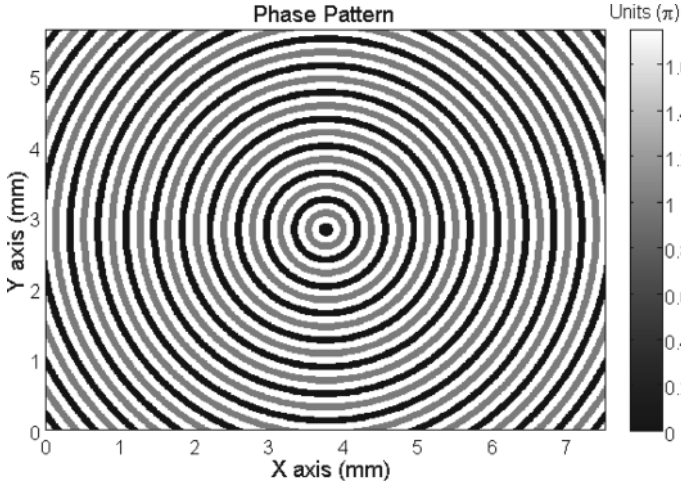


Fig. 2.13 The phase pattern loaded on the SLM. In the radial direction, four rings form a period. Within the period, the rings are loaded with the quadratic phase terms $\exp(-jp^2\pi/N)$, with $p = 0, 1, 2$, and 3 as the order of the ring within the period. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

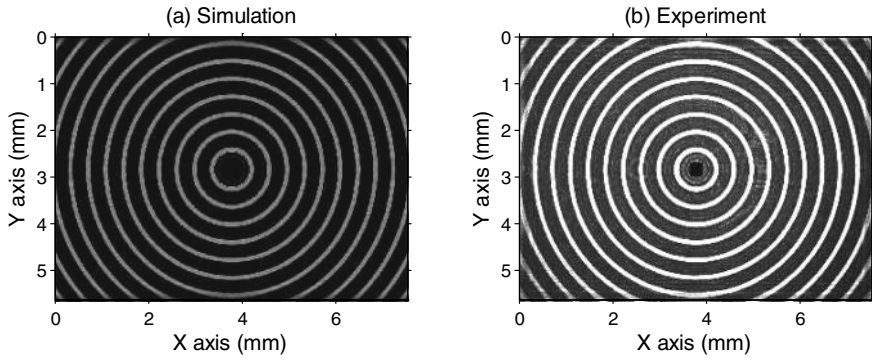


Fig. 2.14 Simulated (a) and experimental (b) results of the Fractional ring shape Talbot effect. The propagation distance is z_{TN} and the topological charge $l = 0$. The center of the rings has been blocked during the illustration to show the rings more clearly. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

this is also referred to as the Talbot carpet. Similarly, we demonstrate Talbot carpet for the ring shape Talbot effect in Figs. 2.15 and 2.16, which clearly shows that the ring-shape Talbot effect on the x - z plane has a periodical behavior. In Fig. 2.15, the evolution of the beam is simulation by the paraxial propagation model while the recorded intensities of 9 frames by the CMOS camera are shown in Fig. 2.16 from z_{TN} to $9z_{TN}$. On each frame, the scale for the x axis is clearly shown, while the scale for the y axis is -0.28 mm to $+0.28$ mm. It should be noted that the propagation

distance in the figure has been normalized with respect to z_{TN} in Fig. 2.15. This suggests the fractional self-imaging effect occurs on the fractional Talbot length and this might offer a way to manipulate the ring shape beam pattern.

Similar to the previous case, the simulated and measured results in Fig. 2.15 and Fig. 2.16 agree well in terms of the positions of the rings but not the amplitude and intensity. The Talbot carpet of the radial Talbot effect suggests that it resembles the lateral Talbot effect very much. The radial Talbot effect also has the fractional Talbot distance at $z = z_{TN}$, and it is clearly shown in the figures. The self-imaging effect is also demonstrated, which can be seen by comparing the results between $z = z_{TN}$ and $z = 5z_{TN}$. These two points are separated by the Talbot length, i.e., $z_{Talbot} = 4z_{TN} = \Lambda^2/\lambda$. As discussed previously, the object will shift half of the period in the radial direction, which is clearly shown in the figure as expected: the intensity distribution at $z = 5z_{TN}$ has the similar intensity distribution as that of the position at $z = z_{TN}$ albeit with a shift by about $\Lambda/2$. Particularly, the center bright spot

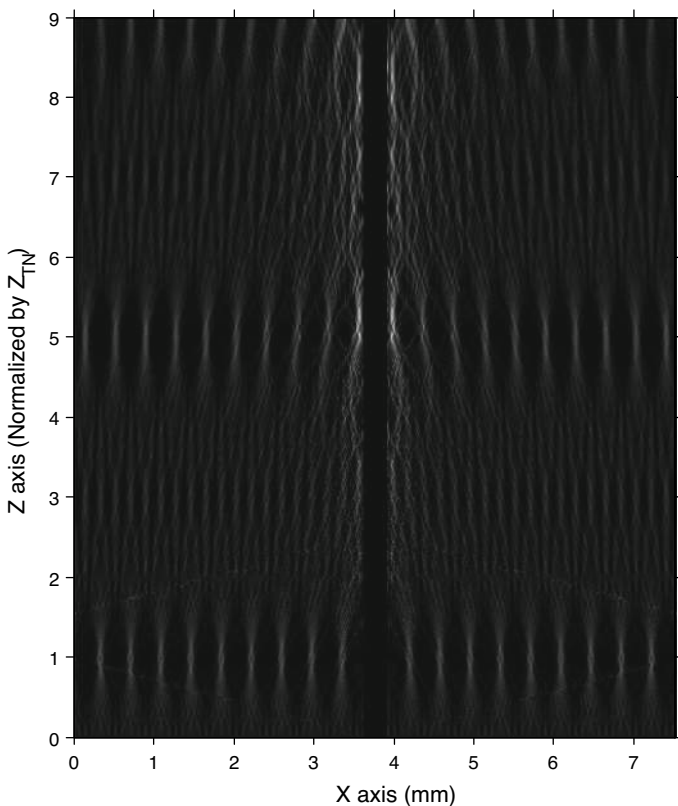


Fig. 2.15 Simulated results of the ring shape Talbot effect on the x - z plane. The propagation distance is 0 to $9z_{TN}$ and the topological charge $l = 0$. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

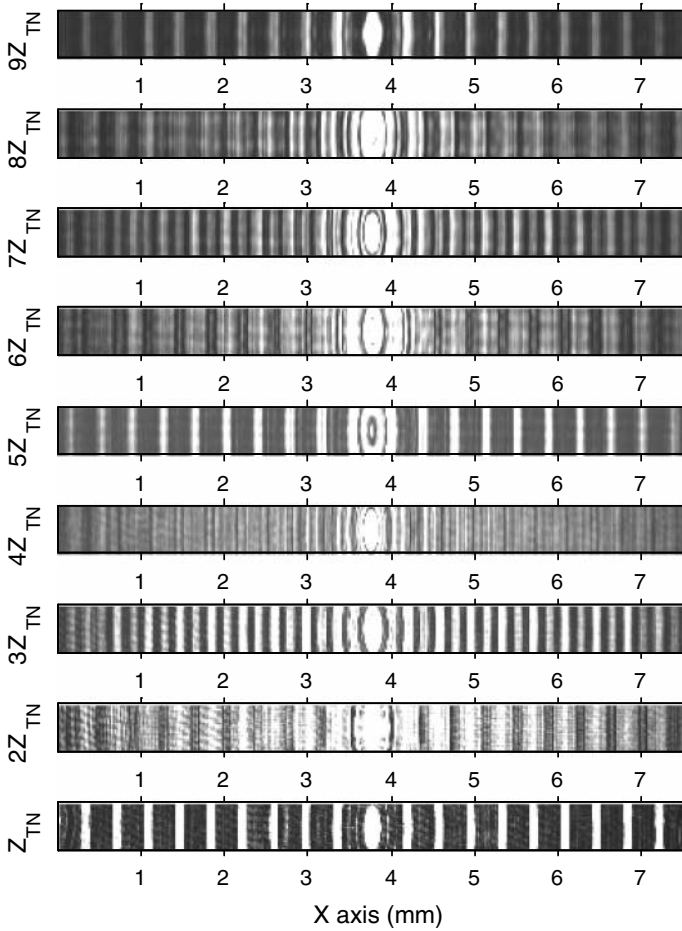


Fig. 2.16 Experimental results of the ring shape Talbot effect. The propagation distance is z_{TN} to $9z_{TN}$ and the topological charge $l = 0$. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

has shifted by $\Lambda/2$ and becomes a ring. At the twice of the standard Talbot length, the field will re-emerge, and this is shown in the figure as well. By comparing the intensity distribution at $z = 9z_{TN}$ and the intensity distribution at $z = z_{TN}$, the two patterns should have identical intensity distribution particularly for the position of the rings. The behaviors are seen more clearly in Fig. 2.17, demonstrating the radial intensity distribution at $z = z_{TN}$, $z = 5z_{TN}$, and $z = 9z_{TN}$. Since the Fig. 2.17 gives the measurement results, the intensity is saturated at the input while the latter two cases show a descending trend in case the radius increases. Such attenuation is caused by the finite rings in the experimental setup.

Other than the fundamental vortex beam, we also tested the theory through simulation and measurement for the beams with other topological charges. The results are

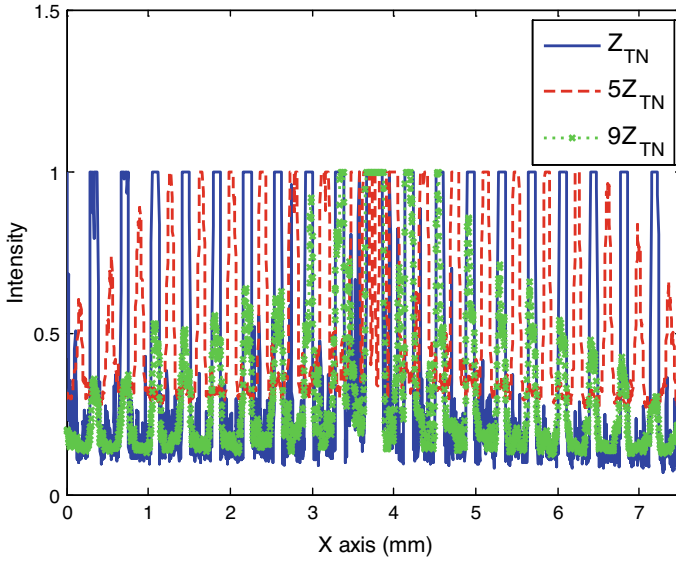


Fig. 2.17 The lateral distribution of the intensity for $z = z_{TN}$, $z = 5z_{TN}$ and $z = 9z_{TN}$. The intensity has been normalized so that the maximum value is 1. The intensity near the center of the rings is strong enough to saturate the CMOS pixels, and they all have the maximum value of 1. Although the intensity at $z = 9z_{TN}$ has significantly attenuated in comparison to the intensity at $z = z_{TN}$, it is quite clear to show that the rings are located in the same positions. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

shown in Fig. 2.18, where the beam possesses the topological charge as 2. Similar to Fig. 2.14, fractional Talbot effect occurred and the two figures agree well in terms of the position of the rings. The Talbot carpet is not shown here, as it resembles the that of the case with topological charge as 0. The simulated phase pattern is shown in Fig. 2.19, which suggests this is a vortex beam with the topological charge of 2.

2.3.4 Summary

In summary, we have reported a novel phenomenon called the radial Talbot effect, which is the results of the interference between the equally spaced ring shape beams. The phenomenon is derived in detailed through the diffraction theory and confirmed by the experimental measurement. The newly discovered effect can be of great potential for the ring shape beam generation, manipulation, and spatial filtering in the cylindrical coordinates to facilitate the applications in communication, biphotonic, atom trapping and laser welding.

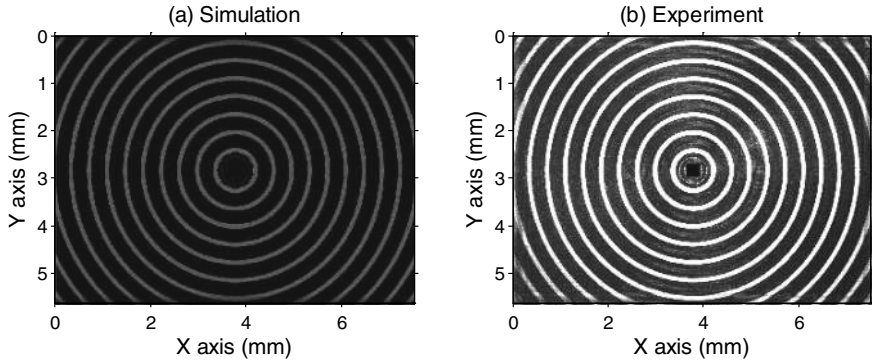


Fig. 2.18 Simulated (a) and experimental (b) results of the Fractional ring shape Talbot effect. The propagation distance is z_{TN} and the topological charge $l = 2$. The center of the rings has been blocked during the illustration to show the rings more clearly. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

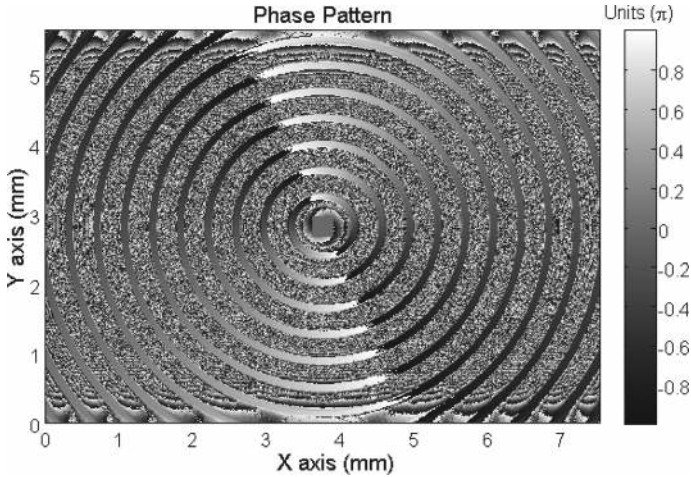


Fig. 2.19 Simulated phase pattern of the Fractional ring shape Talbot effect corresponds to Fig. 2.18. The propagation distance is z_{TN} and the topological charge $l = 2$. The figure is reprinted/adapted with permission from [Phys. Rev. A 103(5), 053,506 (2021)] © The American Physical Society [30]

References

1. J. Zhou, All-optical discrete Fourier transform based on multimode interference couplers. *IEEE Photonics Technol. Lett.* **22**(15), 1093–1095 (2010)
2. W. Shieh, PMD-supported coherent optical OFDM systems. *IEEE Photonics Technol. Lett.* **19**(3), 134–136 (2007)
3. S.L. Jansen, I. Morita, N. Takeda, H. Tanaka, 20-Gb/s OFDM transmission over 4,160-km SSMF enabled by RF-Pilot tone phase noise compensation,” in *Optical Fiber Communication*

- Conference and Exposition and The National Fiber Optic Engineers Conference. OSA Technical Digest Series (CD)* (Optica Publishing Group, Anaheim, California, 2007), p. PDP15
4. W. Shieh, H. Bao, Y. Tang, Coherent optical OFDM: theory and design. *Opt. Express* **16**(2), 841–859 (2008)
 5. H. Sanjoh, E. Yamada, Y. Yoshikuni, Optical orthogonal frequency division multiplexing using frequency/time domain filtering for high spectral efficiency up to 1 bit/s/Hz, in *Optical Fiber Communication Conference and Exhibit* (2002), pp. 401–402
 6. M.E. Marhic, Discrete Fourier transforms by single-mode star networks. *Opt. Lett.* **12**(1), 63–65 (1987)
 7. T. Koichi, O. Manabu, S. Tomohiro, T. Hiroshi, Optical OFDM demultiplexer using silica PLC based optical FFT circuit, in *2009 Conference on Optical Fiber Communication* (2009), pp. 1–3
 8. L.B. Soldano, E.C.M. Pennings, Optical multi-mode interference devices based on self-imaging: principles and applications. *J. Lightwave Technol.* **13**(4), 615–627 (1995)
 9. A.R. Gupta, K. Tsutsumi, J. Nakayama, Synthesis of Hadamard transformers by use of multimode interference optical waveguides. *Appl. Opt.* **42**(15), 2730–2738 (2003)
 10. S.-Y. Tseng, Y. Kim, C.J.K. Richardson, J. Goldhar, Implementation of discrete unitary transformations by multimode waveguide holograms. *Appl. Opt.* **45**(20), 4864–4872 (2006)
 11. J. Zhou, M. Zhang, All-optical discrete sine transform and discrete cosine transform based on multimode interference couplers. *IEEE Photonics Technol. Lett.* **22**(5), 317–319 (2010)
 12. J.M. Heaton, R.M. Jenkins, General matrix theory of self-imaging in multimode interference (MMI) couplers. *IEEE Photonics Technol. Lett.* **11**(2), 212–214 (1999)
 13. W. Zhongde, Fast algorithms for the discrete W transform and for the discrete Fourier transform. *IEEE Trans. Acoust. Speech Signal Process.* **32**(4), 803–816 (1984)
 14. J.M. Heaton, R.M. Jenkins, D.R. Wight, J.T. Parker, J.C.H. Birbeck, K.P. Hilton, Novel 1-to-N way integrated optical beam splitters using symmetric mode mixing in GaAs/AlGaAs multimode waveguides. *Appl. Phys. Lett.* **61**(15), 1754–1756 (1992)
 15. D. Hillerkuss et al., 26 Tbit s⁻¹ line-rate super-channel transmission utilizing all-optical fast Fourier transform processing. *Nat. Photonics* **5**(6), 364–371 (2011)
 16. N.-C. Hu, C.-H. Su, Discrete Fourier preprocessing transforms for the binary phase-only filter. *Appl. Opt.* **34**(8), 1372–1378 (1995)
 17. X. Wu, D. Zou, Z. Dong, X. Zhao, Y. Chen, F. Li, LDPC-coded DFT-spread DMT signal transmission employing probabilistic shaping 16/32QAM for optical interconnection. *Opt. Express* **27**(7), 9821–9828 (2019)
 18. J.W. Cooley, J.W. Tukey, An algorithm for the machine calculation of complex Fourier series. *Math. Comput.* **19**, 297–301 (1965)
 19. N. André, H. Louchet, V. Filsinger, E. Hansen, A. Richter, OFDM and PAM comparison using a high baudrate low resolution IM/DD interface for 400G Ethernet access. *Opt. Express* **24**, 11926 (2016)
 20. I. Kang et al., All-optical OFDM transmission of 7 x 5-Gb/s data over 84-km standard single-mode fiber without dispersion compensation and time gating using a photonic-integrated optical DFT device. *Opt. Express* **19**(10), 9111–9117 (2011)
 21. K. Takiguchi, T. Kitoh, M. Oguma, Y. Hashizume, H. Takahashi, Integrated-optic OFDM demultiplexer using multi-mode interference coupler-based optical DFT circuit, in *Optical Fiber Communication Conference. OSA Technical Digest* (Optica Publishing Group, Los Angeles, California, 2012), p. OM3J.6
 22. J. Zhou, All-optical discrete Fourier transform based on multilevel phase modulation and a dispersive element. *IEEE Photonics Technol. Lett.* **24**(8), 685–687 (2012)
 23. H.F. Talbot, Facts relating to optical science. *Philos. Mag.* **9**, 401–407 (1836)
 24. J. Wen, Y. Zhang, M. Xiao, The Talbot effect: recent advances in classical optics, nonlinear optics, and quantum optics. *Adv. Opt. Photon.* **5**(1), 83–130 (2013)
 25. J. Azaña, H. Guillet de Chatellus, Angular Talbot effect. *Phys. Rev. Lett.* **112**(21), 213902 (2014)
 26. P. Samadian, T. Hall, Cylindrical Talbot effect for ultra-compact multimode interference couplers. *Opt. Lett.* **41**(17), 4110–4113 (2016)

27. D. Hebri, M. Bagheri, S. Rasouli, Talbot effect of azimuthally periodic Bessel-based structures. *Opt. Lett.* **44**(17), 4355–4358 (2019)
28. H. Gao et al., Conformally mapped multifunctional acoustic metamaterial lens for spectral sound guiding and Talbot effect. *Research* **2019** (2019)
29. R.W. Gerchberg, A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik* **35**, 237–246 (1972)
30. J.H. Zhou, J.J. Wu, H.Q. Pu, Ring-shaped Talbot effect. *Phys. Rev. A* **103**(5), 053506 (2021)
31. A. Peruzzo, A. Laing, A. Politi, T. Rudolph, J.L. O'Brien, Multimode quantum interference of photons in multiport integrated devices. *Nat. Commun.* **2**(1), 224 (2011)
32. N.H. Wan, F. Meng, T. Schröder, R.-J. Shiue, E.H. Chen, D. Englund, High-resolution optical spectroscopy using multimode interference in a compact tapered fibre. *Nat. Commun.* **6**(1), 7762 (2015)
33. K. Vynck et al., Ultrafast perturbation maps as a quantitative tool for testing of multi-port photonic devices. *Nat. Commun.* **9**(1), 2246 (2018)
34. B. Santra, C. Baals, R. Labouvie, A.B. Bhattacharjee, A. Pelster, H. Ott, Measuring finite-range phase coherence in an optical lattice using Talbot interferometry. *Nat. Commun.* **8**(1), 15601 (2017)
35. Y. Zhai, C.H. Carson, V.A. Henderson, P.F. Griffin, E. Riis, A.S. Arnold, Talbot-enhanced, maximum-visibility imaging of condensate interference. *Optica* **5**(1), 80–85 (2018)
36. Z. Deng, N. Shapira, R. Remez, Y. Li, A. Arie, Talbot effect in waveforms containing subwavelength multilobe superoscillations. *Opt. Lett.* **45**(9), 2538–2541 (2020)
37. R. Ezaki, Y. Mizutani, N. Ura, T. Uenohara, Y. Makiura, Y. Takaya, Fabrication of three-dimensional high-aspect-ratio structures by oblique-incidence Talbot lithography. *Opt. Express* **28**, 36924 (2020)
38. L. Romero Cortés, H. Guillet de Chatellus, A. Deville, M. Seghilani, I. Hamam, J. Azaña, Generalized Talbot self-healing and noise mitigation of faulty periodic images. *J. Opt. Soc. Am. A* **37**(3), 384–390 (2020)
39. S. Schwarz et al., Talbot effect of orbital angular momentum lattices with single photons. *Phys. Rev. A* **101**(4), 043815 (2020)
40. R. Iwanow, D.A. May-Arrioja, D.N. Christodoulides, G.I. Stegeman, Y. Min, W. Sohler, Discrete Talbot effect in waveguide arrays. *Phys. Rev. Lett.* **95**(5) (2005)
41. S. Rasouli, A.M. Khazaei, D. Hebri, Talbot carpet at the transverse plane produced in the diffraction of plane wave from amplitude radial gratings. *J. Opt. Soc. Am. A* **35**(1), 55–64 (2018)
42. J. Xin, C. Gao, Y. Liu, C. Li, K. Dai, Q. Na, Generation of Bessel beams from a diffractive ring lens. *Opt. Commun.* **310**, 25–30 (2014)
43. V. Garcés-Chávez, D. McGloin, H. Melville, W. Sibbett, K. Dholakia, Simultaneous micro-manipulation in multiple planes using a self-reconstructing light beam. *Nature* **419**(6903), 145–147 (2002)
44. T. Čižmár et al., Generation of multiple Bessel beams for a biophotonics workstation. *Opt. Express* **16**(18), 14024–14035 (2008)
45. W. Wang, Z. Wu, Q. Shang, B. Lu, Propagation of multiple Bessel Gaussian beams through weak turbulence. *Opt. Express* **27**(9), 12780–12793 (2019)
46. R. Sahin, T. Ersoy, S. Akturk, Ablation of metal thin films using femtosecond laser Bessel vortex beams. *Appl. Phys. A* **118**(1), 125–129 (2015)
47. I. Manek, Y.B. Ovchinnikov, R. Grimm, Generation of a hollow laser beam for atom trapping using an axicon. *Optics Commun.* **147**(1), 67–70 (1998)

Chapter 3

Coupled Mode Theory for Multi-mode Propagation



Abstract Coupled mode theory for multi-mode propagation. The chapter introduces the coupled mode theory, which is a powerful tool for multi-mode propagation analysis. The coupled mode theory will be formulated both for the classic optics and quantum optics. While the coupled mode theory is a powerful tool for multi-mode propagation analysis in classic optics, quantum formulation of the multimode propagation is provided afterwards.

In this chapter, the coupled mode theory will be formulated both for the classical optics and quantum optics. While the coupled mode theory is a powerful tool for multi-mode propagation analysis in classical optics, quantum formulation of the multimode propagation is provided afterwards.

3.1 Coupled Mode Theory in Classical Optics

Part of the section reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [1].

3.1.1 Coupled Mode Theory Inside Optical Waveguides with the Scalar Wave Equation

To start with, we introduce the scalar Helmholtz equation [2–5], which can serve as the basis for further derivation,

$$\nabla^2 \varphi + k^2 \varphi = 0 \quad (3.1)$$

where k stands for the wave number in the waveguides. Although it is in the scalar form, it can be easily extended to the vector form [2] in the latter section.

The wave propagation is impacted the random index variations which can be caused by the fabrication process in the waveguides or the weather change in free space. After considering the index perturbation during the propagation, we have

$$\begin{aligned}\nabla^2\varphi + \delta k^2\varphi + \langle k^2 \rangle\varphi &= 0 \\ \delta k^2 &= k^2 - \langle k^2 \rangle\end{aligned}\quad (3.2)$$

where $\langle \rangle$ stands for the average value. It is known that the field, which exists either in the optical waveguide or the free space can be expanded by the orthonormal basis. In the optical waveguides, these orthonormal functions can be the optical waveguide modes [2]

$$\varphi = \sum a_m \phi_m e^{-j\beta_m z} \quad (3.3)$$

where, a_m stands for the amplitude, β_m the propagation constant, and ϕ_m the modal field of the modes. Here, Σ not only stands for the summation for the guided modes, but also the integration for the leaky modes which have continuous varying propagation constants. These orthonormal bases, i.e., the orthogonal modes inside the waveguides should fulfill the following equation in the waveguides without index perturbation:

$$\nabla_{\perp}^2 \phi_m + (\langle k^2 \rangle - \beta_m^2) \phi_m = 0 \quad (3.4)$$

Substituting Eqs. (3.3–3.4) into Eq. (3.2), and using the paraxial approximation, one has [2]

$$-\sum \phi_m 2j\beta_m \frac{\partial a_m(z)}{\partial z} e^{-j\beta_m z} + \delta k^2 \sum a_{m'}(z) \phi_{m'} e^{-j\beta_{m'} z} = 0 \quad (3.5)$$

Multiplying the complex conjugate of the mode field on Eq. (3.5) and integrate over the cross section of the waveguide, we have [2]

$$\frac{da_m(z)}{dz} = \sum C_{mm'}(z) e^{j(\beta_m - \beta_{m'})z} a_{m'}(z) \quad (3.6)$$

where

$$C_{mm'}(z) = \frac{1}{2j\beta_m} \frac{\iint_{xy} \phi_m^* \phi_{m'} \delta k^2 dS}{\iint_{xy} |\phi_m|^2 dS} \quad (3.7)$$

The modes are usually normalized as [2–5]

$$\iint_{xy} \phi_m \phi_n^* dS = \frac{2\omega\mu}{\beta_m} \delta_{mn} \quad (3.8)$$

And therefore, the coefficient can be evaluated as:

$$\begin{aligned} C_{mm'}(z) &= \frac{k_0^2}{j4\omega\mu} \iint_{xy} \phi_m^* \phi_{m'} \delta n^2 dS \\ \delta k^2 &= \delta n^2 k_0^2 \\ \delta n^2 &= n^2 - \langle n^2 \rangle \end{aligned} \quad (3.9)$$

where k_0 is the wave number in the free space.

3.1.2 Extension of Formulas in the Vector Theory

The formulas in the previous section have been presented in the scalar form while the more general case of the vector form formulas is required for the strongly guided waveguides. Similar coupled mode equations in the form of Eq. (3.6) can be derived. It is, however, with the coupling coefficient to be evaluated by the vector mode field [6]

$$C_{mm'}(z) = \frac{-j\omega\epsilon_0 \iint_{xy} \mathbf{e}_m^* \cdot \mathbf{e}_{m'} \delta n^2 dS}{\iint ((\mathbf{e}_m \times \mathbf{h}_m^*) + (\mathbf{e}_m^* \times \mathbf{h}_m)) \cdot \vec{z} dS} \quad (3.10)$$

where $\mathbf{e}_m$ and $\mathbf{h}_m$ denote the electric field and the magnetizing field of the m th modes. If we define [6]

$$\left(P_m = \frac{1}{4} \iint ((\mathbf{e}_m \times \mathbf{h}_m^*) + (\mathbf{e}_m^* \times \mathbf{h}_m)) \cdot \vec{z} dS = 1 \right) \quad (3.11)$$

we have

$$C_{mm'}(z) = \frac{k_0^2}{j4\omega\mu_0} \iint_{xy} \mathbf{e}_m^* \cdot \mathbf{e}_{m'} \delta n^2 dS \quad (3.12)$$

which is almost in the same form as the scalar case. During the derivation of Eq. (3.10), the orthogonality between the modes is used [7]. The readers can find the derivations in [8] and in Marcuse' section [9]. Hence, the vector form and the scalar form coupled mode theories share the same expressions, albeit with different definitions for the coupling coefficients.

3.2 Characterization of PT-Symmetric Quantum Interference Based on the Coupled Mode Theory

Part of the section reprinted/adapted with permission from [Opt. Express 30, 23,600–23,607 (2022)] © The Optical Society of America [10].

In the previous section, the coupled mode theory for the classical optics is presented. It is, however, highly expected to be extended for the quantum optics. In this section, we extend the coupled mode theory in the quantum theoretical framework, which can be used to investigate the quantum interference inside the optical waveguide system. A few important quantum optics problems, such as the parity-time (PT) symmetric property for the two-waveguide system can be studied thoroughly. Such a problem was only formulated for the waveguide system with equal losses, which is extended in this section for the case with two optical waveguides with unequal losses. The theory can be used to study the entangled state with two photons. The expression for the well-known Hong-Ou-Mandel (HOM) dip is derived, which is in an exact agreement with the published results. A special one-photon quantum interference phenomenon is predicted, which has not been investigated before. A quantum interference process similar to the HOM effect can be expected with only one photon emerged at the output of the two-waveguide system while the other photon is absorbed due to loss. The phenomenon is unique in the PT symmetric system as it is not possible to exist in the equal loss system.

3.2.1 Introduction

Optical waveguides have wide applications, such as information transmission, light matter interaction, optical beam manipulation, etc. Various photonics passive and active functional devices have been fabricated based on the waveguide structure, and among them, the devices to manipulate the quantum states of photons have received special attentions [11–17]. While the optical wave propagation inside the waveguides is studied thoroughly by the classical optics [18], the quantum interactions should be formulated in the quantum optics framework [19–23]. Loss in the system is usually difficult to deal with in the quantum framework. The traditional quantum theories treat the loss in the single mode waveguides as the coupling between the propagation mode and the modes in the reservoir, it is, however, not extended to the multi-waveguide/multimode case. Luks et al. [20], Liñares et al. [21, 22] and Barral [23] proposed the multimode waveguide quantum wave propagation theory, which includes the constant coupling between the different modes or the linear and nonlinear losses in the waveguides [20, 23]. However, simultaneous consideration of the coupling and the unequal losses were not included. In [24], two waveguides with equal attenuations were studied in the quantum framework. It is, however, only a special extension of the none-loss case because the solution remains the same despite some lossy terms added, and the Hermitian property system is not affected.

Recently, it is found that quantum interaction inside the coupled waveguides with unequal losses could result in the parity time (PT) symmetric property [15], which is a unique feature possessing the real Eigen values other than Hermitian systems [25]. PT symmetry can exist both in quantum optics [15, 16] and classical optics [26–32]. For the PT symmetry problem in classical optics, the well-known coupled mode theory can be applied [26–32], while the quantum coupled mode theory remains unmatre as explained in the previous paragraph and quite complicated formulation is usually required. For example, a Lie algebra approach was suggested in [15]. In addition to its tedious mathematical presentation, it is difficult to be extended for the cases with more than two coupled waveguides. The obscure mathematical language in [15] also makes it difficult to bring a clear physical picture for the entangled state. For example, if two photons are injected into the two-waveguide system with unequal losses and one photon is lost due to the attenuation, will quantum entangled state be preserved for the single photon? If so, will the quantum entangled state be detected through the method called the Hong-Ou-Mandel (HOM) effect [15] these problems remain unanswered.

In this section, a coupled mode theory approach with the introduction of the Langevin noise is proposed to offer a flexible tool to formulate the quantum interference inside the coupled optical waveguides with unequal losses. It is well-known that the interference of two indistinguishable photons can be detected through the measurement of the Hong-Ou-Mandel (HOM) effect. The expression for the HOM dip, which has been presented in the published literature [15], can be obtained analytically by the proposed theory, which verify the correctness of the theory.

In addition to that, the one photon entangled quantum state is formulated for the first time. The entangled state is caused by the interference for the photon pair with one of the photons absorbed by the waveguide attenuation in the PT symmetric systems. It is unclear which photon is lost until the detection takes places. It is expected that a dip/peak similar to the HOM dip/peak [33, 34] can be detected at the receiver side, which can never be expected in a Hermitian system. The one photon entanglement is associated with the Langevin noise terms induced by the unequal losses, which makes the noise term to be correlated and interactions to occur.

3.2.2 *Coupled Mode Theory for the Coupled Waveguides with Unequal Losses*

The Hamiltonian operator in two coupled optical waveguides under the interaction picture can be formulated by [15]:

$$M = \hbar \left(\kappa \hat{a}_1^\dagger \hat{a}_2 + \kappa \hat{a}_2^\dagger \hat{a}_1 \right), \quad (3.13)$$

where $\hbar$ is the plank constant, κ is the coupling coefficient between the two waveguides, $\hat{a}_m$ and $\hat{a}_m^\dagger$ are the annihilation and creation operators of the two waveguides,

with m indicating the waveguide number. Under the Heisenberg picture and in the absence of waveguide loss, we have the following coupled mode equation [15],

$$\frac{\partial}{\partial z} \hat{a}_m = \frac{1}{i\hbar} [\hat{a}_m, M]. \quad (3.14)$$

Equation (3.14) is the same as the coupled mode equation derived in the classical optics [18]

$$\begin{aligned} \frac{d\hat{a}_1(z)}{dz} &= -i\kappa\hat{a}_2(z), \\ \frac{d\hat{a}_2(z)}{dz} &= -i\kappa\hat{a}_1(z). \end{aligned} \quad (3.15)$$

While the quantum coupled mode theory seems simple and straight forward, it differs remarkably from its classical optics counterpart if the loss terms are introduced [19, 35]

$$\begin{aligned} \frac{d\hat{a}_1(z)}{dz} &= -i\kappa\hat{a}_2(z) - \gamma_1\hat{a}_1(z) + \hat{f}_1(z), \\ \frac{d\hat{a}_2(z)}{dz} &= -i\kappa\hat{a}_1(z) - \gamma_2\hat{a}_1(z) + \hat{f}_2(z), \\ [\hat{f}_m(z), \hat{f}_{m'}^\dagger(z')] &= 2\delta_{mm'}\gamma_m\delta(z - z'), \\ \langle \hat{f}_m^\dagger(z)f_m(z') \rangle &= 2\bar{n}_\sigma\gamma_m\delta(z - z'), \\ \langle f_m(z)f_m^\dagger(z') \rangle &= 2(1 + \bar{n}_\sigma)\gamma_m\delta(z - z'), \end{aligned} \quad (3.16)$$

where γ_m is the loss term on the m th waveguide, f_m is Langevin noise term respectively, $\bar{n}_\sigma$ is the mode density [19, 35]:

$$\bar{n}_\sigma = \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \approx 0. \quad (3.17)$$

which is close to zero at the room temperature [35]. The most significant difference between the quantum coupled mode theory and the classic coupled mode theory is the noise term, which is introduced to ensure the commutation relation for the operators [19, 35]. As will be shown later, the noise terms will play important roles in the quantum interactions. Equation (3.16) can be solved in the close form:

$$\begin{pmatrix} \hat{a}_1(z) \\ \hat{a}_2(z) \end{pmatrix} = \exp\left(-\left(\frac{\gamma_1 + \gamma_2}{2}\right)z\right) \begin{pmatrix} \cos(\lambda z) + \frac{\gamma_2}{\lambda} \sin(\lambda z) & -\frac{i\kappa}{\lambda} \sin(\lambda z) \\ -\frac{i\kappa}{\lambda} \sin(\lambda z) & \cos(\lambda z) - \frac{\gamma_1}{\lambda} \sin(\lambda z) \end{pmatrix}$$

$$\begin{pmatrix} \hat{a}_1(0) \\ \hat{a}_2(0) \end{pmatrix} + F(z), \quad (3.18)$$

with:

$$\begin{aligned} \gamma &= \frac{\gamma_2 - \gamma_1}{2}, \\ \lambda &= \sqrt{\kappa^2 - \gamma^2}, \\ F(z) &= \begin{pmatrix} F_1(z) \\ F_2(z) \end{pmatrix} = \exp\left(-\left(\frac{\gamma_1 + \gamma_2}{2}\right)z\right) \\ &\quad \times \int_0^z \exp\left(\left(\frac{\gamma_1 + \gamma_2}{2}\right)s\right) \\ &\quad \begin{pmatrix} \cos(\lambda(z-s)) + \frac{\gamma}{\lambda} \sin(\lambda(z-s)) & -\frac{i\kappa}{\lambda} \sin(\lambda(z-s)) \\ -\frac{i\kappa}{\lambda} \sin(\lambda(z-s)) & \cos(\lambda(z-s)) - \frac{\gamma}{\lambda} \sin(\lambda(z-s)) \end{pmatrix} \\ &\quad \begin{pmatrix} \hat{f}_1(s) \\ \hat{f}_2(s) \end{pmatrix} ds. \end{aligned} \quad (3.19)$$

Based on the above equation, one may evaluate the noise term correlation matrix analytically as:

$$\begin{aligned} \langle F(z)F^H(z) \rangle &= \begin{pmatrix} \langle |F_1(z)|^2 \rangle & \langle F_1(z)F_2^*(z) \rangle \\ \langle F_1(z)^*F_2(z) \rangle & \langle |F_2(z)|^2 \rangle \end{pmatrix} \\ &= 2 \int_0^z \exp(-(\gamma_1 + \gamma_2)s) A(s) \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} A^H(s) ds, \\ A(s) &= \begin{pmatrix} \cos(\lambda s) + \frac{\gamma}{\lambda} \sin(\lambda s) & -\frac{i\kappa}{\lambda} \sin(\lambda s) \\ -\frac{i\kappa}{\lambda} \sin(\lambda s) & \cos(\lambda s) - \frac{\gamma}{\lambda} \sin(\lambda s) \end{pmatrix}. \end{aligned} \quad (3.20)$$

One may draw a conclusion from the above equation that the correlation matrix has the non-diagonal terms to be zero if the losses are equal, which suggests the two noise terms are not correlated. However, the non-diagonal terms become non-zero if the losses become unequal, which suggests the noise terms are correlated.

3.2.3 Analysis of PT Symmetric Quantum Entanglement

Considering that two indistinguishable photons are injected into two waveguides and create the state $|1, 1\rangle$,

$$|1, 1\rangle = \hat{a}_1^\dagger \hat{a}_2^\dagger |0, 0\rangle. \quad (3.21)$$

The creation operation will evolve along the propagation distance according to the coupled mode equation, and therefore, we have:

$$\begin{aligned} \hat{a}_1^\dagger(z)\hat{a}_2^\dagger(z)|0, 0\rangle = & \beta_{|2,0\rangle}|2, 0\rangle + \beta_{|0,2\rangle}|0, 2\rangle + \beta_{|1,1\rangle}|1, 1\rangle \\ & + \beta_{|1,0\rangle}|1, 0\rangle + \beta_{|0,1\rangle}|0, 1\rangle + \beta_{|0,0\rangle}|0, 0\rangle, \end{aligned} \quad (3.22)$$

with the coefficients as:

$$\begin{aligned} \beta_{|2,0\rangle} &= \sqrt{2} \exp(-(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right) \frac{i\kappa}{\lambda} \sin(\lambda z), \\ \beta_{|0,2\rangle} &= \sqrt{2} \exp(-(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right) \frac{i\kappa}{\lambda} \sin(\lambda z), \\ \beta_{|1,1\rangle} &= \exp(-(\gamma_1 + \gamma_2)z) \left(\left(\cos(\lambda z)^2 - \frac{\gamma^2}{\lambda^2} \sin(\lambda z)^2 \right) - \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \right), \\ \beta_{|1,0\rangle} &= \exp\left(-\left(\frac{\gamma_1 + \gamma_2}{2}\right)z\right) \left(\left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right) F_2^\dagger + \frac{i\kappa}{\lambda} \sin(\lambda z) F_1^\dagger \right), \\ \beta_{|0,1\rangle} &= \exp\left(-\left(\frac{\gamma_1 + \gamma_2}{2}\right)z\right) \left(\left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right) F_1^\dagger + \frac{i\kappa}{\lambda} \sin(\lambda z) F_2^\dagger \right), \\ \beta_{|0,0\rangle} &= F_1^\dagger F_2^\dagger. \end{aligned} \quad (3.23)$$

The probabilities to detect the related quantum states are:

$$\begin{aligned} p_{|2,0\rangle} &= 2 \exp(-2(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2, \\ p_{|0,2\rangle} &= 2 \exp(-2(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2, \\ p_{|1,1\rangle} &= \exp(-2(\gamma_1 + \gamma_2)z) \left(\frac{\kappa^2 \cos(2\lambda z) - \gamma^2}{\lambda^2} \right)^2, \\ p_{|1,0\rangle} &= \exp(-(\gamma_1 + \gamma_2)z) \\ &\quad \times \left(\left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \langle F_2 F_2^\dagger \rangle + \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \langle F_1 F_1^\dagger \rangle \right. \\ &\quad \left. + 2\text{Re} \left(- \left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right) \frac{i\kappa}{\lambda} \sin(\lambda z) \langle F_1 F_2^\dagger \rangle \right) \right), \\ p_{|0,1\rangle} &= \exp(-(\gamma_1 + \gamma_2)z) \\ &\quad \times \left(\frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \langle F_2 F_2^\dagger \rangle + \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \langle F_1 F_1^\dagger \rangle \right. \\ &\quad \left. + 2\text{Re} \left(- \frac{i\kappa}{\lambda} \sin(\lambda z) \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right) \langle F_2 F_1^\dagger \rangle \right) \right), \\ p_{|0,0\rangle} &= \left(\langle F_1 F_1^\dagger \rangle \langle F_2 F_2^\dagger \rangle + \langle F_2 F_1^\dagger \rangle \langle F_1 F_2^\dagger \rangle \right). \end{aligned} \quad (3.24)$$

The HOM dip is defined as the probability to detect two photons at the both waveguides, and can be calculated as

$$p_{|1,1\rangle} = \exp(-2(\gamma_1 + \gamma_2)z) \left(\frac{\kappa^2 \cos(2\lambda z) - \gamma^2}{\lambda^2} \right)^2, \quad (3.25)$$

which is exactly the same as the results derived by the complex mathematics [15]. The important parameter visibility is defined as $\frac{p_{|1,1\rangle}}{p_{|1,1\rangle d}} - 1$, where $p_{|1,1\rangle d}$ stands for the probability in case the two photons are distinguishable, e.g. two photons with different polarizations. It is quite straight forward to provide the expressions for the evolution of the distinguishable photon pairs:

$$\begin{aligned} p_{|2,0\rangle d} &= 2 \exp(-2(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2, \\ p_{|0,2\rangle d} &= 2 \exp(-2(\gamma_1 + \gamma_2)z) \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2, \\ p_{|1,1\rangle d} &= \exp(-2(\gamma_1 + \gamma_2)z) \left(\left(\cos(\lambda z)^2 - \frac{\gamma^2}{\lambda^2} \sin(\lambda z)^2 \right)^2 + \left(\frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \right)^2 \right), \\ p_{|1,0\rangle d} &= \exp(-(\gamma_1 + \gamma_2)z) \\ &\quad \left(\left(\cos(\lambda z) + \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \langle F_2 F_2^\dagger \rangle + \frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \langle F_1 F_1^\dagger \rangle \right), \\ p_{|0,1\rangle d} &= \exp(-(\gamma_1 + \gamma_2)z) \\ &\quad \left(\frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \langle F_2 F_2^\dagger \rangle + \left(\cos(\lambda z) - \frac{\gamma}{\lambda} \sin(\lambda z) \right)^2 \langle F_1 F_1^\dagger \rangle \right), \\ p_{|0,0\rangle d} &= \langle F_1 F_1^\dagger \rangle \langle F_2 F_2^\dagger \rangle. \end{aligned} \quad (3.26)$$

Therefore, we may calculate $\frac{p_{|1,1\rangle}}{p_{|1,1\rangle d}} - 1$ as

$$\frac{-2 \left(\cos(\lambda z)^2 - \frac{\gamma^2}{\lambda^2} \sin(\lambda z)^2 \right) \left(\frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \right)}{\left(\cos(\lambda z)^2 - \frac{\gamma^2}{\lambda^2} \sin(\lambda z)^2 \right)^2 + \left(\frac{\kappa^2}{\lambda^2} \sin(\lambda z)^2 \right)^2}. \quad (3.27)$$

To investigate the correctness of the proposed theory, a case presented in the published literature is studied. An indistinguishable photon pair is injected into the two-waveguide system, which has the coupling strength $\kappa = 0.25 \text{ cm}^{-1}$ and the unequal losses $\gamma_1 = 0 \text{ cm}^{-1}$ and $\gamma_2 = 0.35 \text{ cm}^{-1}$ [15].

As is usually done for the HOM effect measurement, the coincidence to observe the two photons simultaneously at the output of the two waveguides is shown as the curves in Fig. 3.1a. In the figure, the blue solid line stands for the coincidence to detect the two indistinguishable photons in the PT symmetric system while the red dashed line

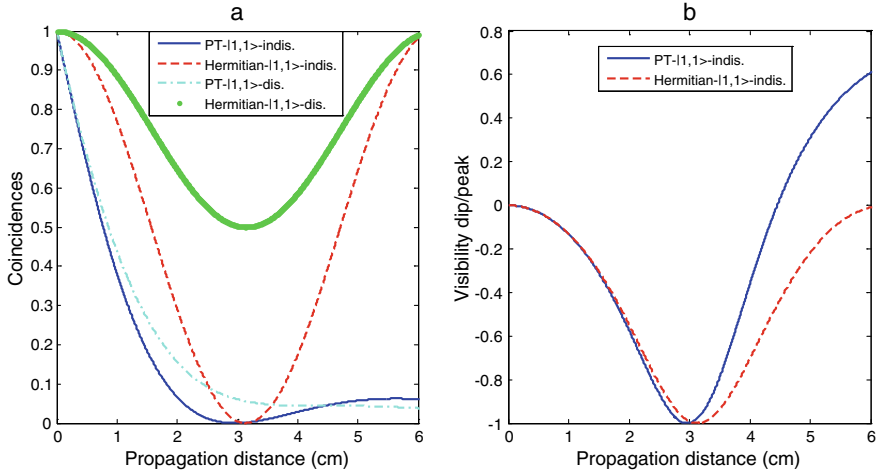


Fig. 3.1 the coincidence and the visibility for the two-photon state varying with respect to propagation distance. Dis.—stands for distinguishable photon pairs and indis.—stands for the indistinguishable photon pairs. The figure is reprinted/adapted with permission from [Opt. Express 30, 23,600–23,607 (2022)] © The Optical Society of America [10]

stands for the coincidence to detect the two indistinguishable photons in the Hermitian systems (with the losses as 0). The cyan dash-dotted line and the green dotted line stand for the coincidence to detect the two distinguishable photons respectively. One may conclude from the figure that the coincidence varies along the propagation distance and the difference between the distinguishable and indistinguishable photons also varies, which leads to the visibility of the HOM effect as is shown in Fig. 3.1b. In Fig. 3.1b, the deepest HOM dip shifts from 3 cm in the Hermitian system to about 2.8 cm in the PT symmetric system. Such HOM dip shift has been observed in [15] with also identical quantity.

After analyzing the two-photon interaction, we report a new phenomenon which has not been studied before, i.e., the one photon interaction, which can be studied thorough the expressions for $p_{|1,0\rangle}$ and $p_{|0,1\rangle}$. These two states are single photon states with one photon loss due to the attenuation. The expressions in Eq. (3.24) suggest that the two probabilities contain the cross correlation of the two Langevin noise terms, which suggests interference may happen. Such interference will not occur if the losses are equal because the correlation term will be zero. The results are clearly demonstrated in Fig. 3.2. Figure 3.2a shows the probability evolution along the propagation distance. The blue solid line and the red dashed line stands for the probabilities of the states $|1, 0\rangle$ and $|0, 1\rangle$ to be detected at the receiver side. The distinguishable photon pair results are plotted as the cyan dash-dotted and green dotted lines. Similar to the HOM effect for the two-photon pairs, we may expect the one-photon HOM effect, which indicate the difference between the distinguishable and indistinguishable photons. Such difference is shown in Fig. 3.2b, which shows that state $|1, 0\rangle$ will have an interference dip and it is like the HOM dip for the

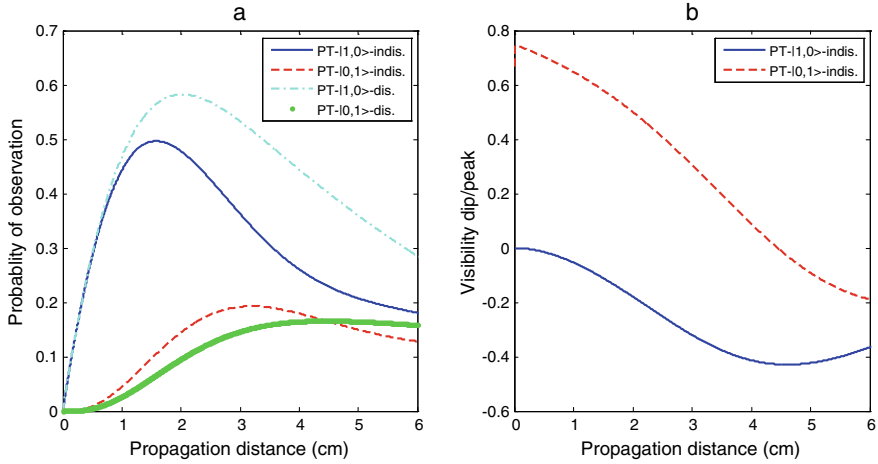


Fig. 3.2 Interference of the one-photon state with the other photon lost due to the waveguide attenuation. Dis.—stands for distinguishable photon pairs and indis.—stands for the indistinguishable photon pairs. State $|1, 0\rangle$ has the interference dip, while state $|0, 1\rangle$ has the interference peak. The figure is reprinted/adapted with permission from [Opt. Express 30, 23,600–23,607 (2022)] © The Optical Society of America [10]

two-photon case. This indicates that the one photon interference can be potentially measured by implementing one photon counter at the waveguide 1 output. Such phenomenon has never been reported or observed before.

The one photon interference phenomenon can be explained by the fact that the one-photon state is actually formed by the two-photon pair at different input of the waveguides. During the propagation inside the two-waveguide system, one of the photons is lost and it is replaced with the Langevin noise photon. However, it is unclear which photon has been absorbed and therefore, two noise terms are added in the probability expression. The noise terms are correlated due to the unequal losses and therefore, they may interfere. In the Hermitian system or in the waveguides with equal losses, the noise term correlation is zero. Therefore, such a phenomenon can not be observed, and this is the reason it has not been reported before.

Other probabilities $p_{|2,0\rangle}, p_{|0,2\rangle}$ and $p_{|0,0\rangle}$ in the expressions are investigated in Fig. 3.3. As can be seen from the figure, the probability for the state of $|0, 0\rangle$ increases when the propagation distance increases, which is the natural result of the waveguide attenuation. It is confirmed that the summation of all the probabilities equals 1, which is shown as the black dot line, i.e.,

$$p_{|2,0\rangle} + p_{|0,2\rangle} + p_{|1,1\rangle} + p_{|1,0\rangle} + p_{|0,1\rangle} + p_{|0,0\rangle} = 1. \quad (3.28)$$

The results suggest that the proposed theory is self-consistent.

Afterwards, a similar case for the waveguides with unequal losses is investigated with the two lossy waveguides having the losses $\gamma_1 = 0.15 \text{ cm}^{-1}$ and $\gamma_2 = 0.5 \text{ cm}^{-1}$.

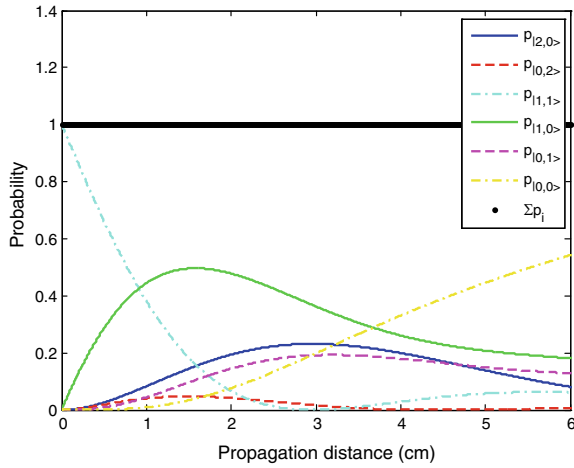


Fig. 3.3 The evolution behaviors of the probabilities $p_{|2,0>}$, $p_{|0,2>}$ and $p_{|0,0>}$ along the propagation distance. The figure is reprinted/adapted with permission from [Opt. Express 30, 23,600–23,607 (2022)] © The Optical Society of America [10]

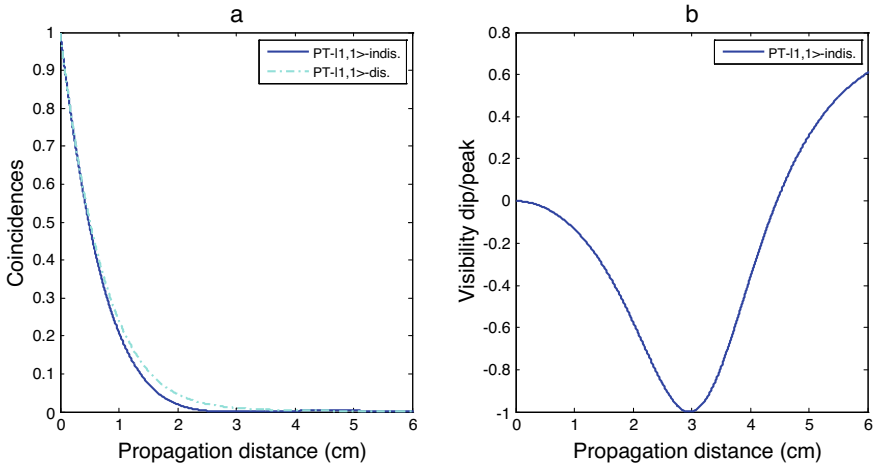


Fig. 3.4 the coincidence and the visibility for the two-photon state varying with respect to propagation distance in two lossy waveguides with $\gamma_1 = 0.15 \text{ cm}^{-1}$ and $\gamma_2 = 0.5 \text{ cm}^{-1}$. Dis.—Stands for distinguishable photon pairs and indis.—stands for the indistinguishable photon pairs. The figure is reprinted/adapted with permission from [Opt. Express 30, 23,600–23,607 (2022)] © The Optical Society of America [10]

The results are shown in Fig. 3.1, which shows identical visibility as the previous case. Because the loss difference between the two waveguides is unchanged and so is the visibility.

3.2.4 Summary

In this section, we have presented a quantum coupled mode theory for the characterization of quantum light propagation inside two coupled waveguides with unequal losses. Such unequal loss will lead to the so-called PT symmetric system. The theory derives the expression for the HOM effect and predicts an unprecedented one-photon interference phenomenon. The proposed theory is studied in detail and compared with the published results with excellent agreement. It can be extended to the multiple waveguide cases with the waveguide number larger than 2, which will be shown in the next section.

3.3 Analytical Formulation of Quantum Interference Inside Coupled Waveguides with Unequal Losses

Part of the section reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36].

In this section, we have proposed a quantum coupled mode theory to study the quantum state evolution inside optical waveguides with unequal losses. The coupled mode equation includes the mode coupling and the Langevin noise terms to ensure the commutation relationship for the operators. Based on the analytical solution of the coupled mode equation, it is possible to find a close form formula for the correlation matrix of the Langevin noise terms. Based on the correlation matrix, it is possible to find the analytical solution for the density matrix for the propagation states. A three-waveguide system and an 89-waveguide system are analyzed by the theory.

3.3.1 Introduction

Photon manipulation has become one of the most heated topics in quantum optics [11–14, 17]. As is shown in the previous section, photon manipulation based on the coupled waveguides are feasible and adjustable [15, 16, 25–32, 37, 38]. For an example, the previously mentioned a parity time (PT) symmetric system was proposed to be realized in a coupled waveguide system [15, 25–32]. While only the two waveguides are taken into consideration in the previous section, the number of coupled waveguides might be increased. For example, three coupled waveguides

could be used to form an anti-PT symmetric system if the middle waveguide has a much higher loss [16, 37]. If the number of the coupled waveguides could be increased beyond three to realize quantum information processing device, e.g., the quantum fractional Fourier transformer [38].

As is discussed in the previous section, the coupled mode theory in classical optics is quite mature and has been studied thoroughly [8, 18]. Anyway, the quantum handling of wave coupling phenomenon is different and complex [19–24, 39–42]. One simple approach is to use the beam splitter model to mimic the photon transferring in a discrete manner [40–42], and the Langevin noise [40] should be added to ensure the commutation relationship. The frequency dependent photon incidence could be considered by this simple model [41, 42]. Due to the discrete nature of the model, it is difficult to model the distributed coupling process which transfers the photon bit by bit. Furthermore, the beam splitter model has only two input ports and two output ports, which suggests that it is not possible to be extended for the case with more than two coupled waveguides/modes. The problem might be solved by employing the coupled mode theory in quantum optics, which includes the coupling among the modes. It is, however, difficult to tackle the mode losses as the losses are usually treated by the reservoir model [19], which is quite complex in the multimode case. As is mentioned in the previous section, Luks et al. [20], Liñares et al. [21, 22] and Barral's thesis [23] have proposed important theoretical models based on the coupled mode theory; however, the problem of loss treatment under the multimode coupling scenario has not been fully solved. In [24] and [15], the coupled waveguide with equal losses and unequal losses were considered. It is, however, only applicable for the case with two waveguides. It is difficult to be extended to the cases with more than three waveguides.

As is mentioned above, the waveguides array with more than three waveguides can be used to realize the PT symmetric [15, 25–32] and anti-PT symmetric systems [16, 37], as well as many other functional devices [38]. There is a remarkable requirement for a more general coupled mode theory which can tackle the coupled waveguides with unequal loss. We have derived a comprehensive theory based on the coupled mode equation with Langevin noise terms, which has been successfully used to tackle the two coupled waveguides with unequal losses [10]. In this section, the theory is extended for the case with the waveguide number beyond 3.

Similar to the previous section, the coupled mode equations can be solved by the simple mathematics and the close form formula for the correlation matrix of the Langevin noise terms can be obtained. Based on the correlation matrix, it is possible to compute the density matrix of the state and to resolve the problem of quantum coupling. A three-waveguide system, which used to mimic an anti-PT symmetric system [16], is simulated and studied in detail, which confirms the assertion in the published literature.

For example, the distinguishable and indistinguishable photon pair will exhibit different interference behaviors and it is called the Hong-Ou-Mandel (HOM) effect, and a coincidence rate dip can be observed for the indistinguishable photons [33, 42]. In the anti-PT symmetric three-waveguide system, such an HOM effect does exist and can be observed, which is confirmed by the simulation by the proposed theory.

The complex entangled state in an 89-waveguide system is simulated afterwards with the density matrix fully evaluated analytically.

3.3.2 Coupled Mode Theory for the Coupled Waveguides with Unequal Losses

To start with, we present the Hamiltonian in a N-waveguide optical system as follows [15]:

$$\left(H = \left\{ \sum_m \hbar \beta_m \hat{a}_m \hat{a}_m^\dagger + \left(\sum_{m < m'} \hbar C_{mm'} \hat{a}_m \hat{a}_{m'}^\dagger + h.c. \right) \right\}, \right) \quad (3.29)$$

where $\hbar$ is the plank constant, β_m is the propagation constant of the m th waveguide, $C_{mm'}$ is the coupling coefficient between the m th and the m' th waveguides, and $\hat{a}_m$ and $\hat{a}_m^\dagger$ are the annihilation and creation operators of the m th waveguide. The coupled mode equation can be derived based on the Hamiltonian in a N-waveguide optical system [15]:

$$\frac{d}{dz} \hat{a}_m = \frac{1}{i\hbar} [\hat{a}_m, H] = -i\beta_m \hat{a}_m - i \sum C_{mm'} \hat{a}_{m'}, \quad (3.30)$$

which is identical to the coupled mode equation in classical optics [8, 18]. As is mentioned in the previous section, the main innovation of the work is to include the unequal loss term in the quantum coupled mode equation [10, 19, 35]:

$$\frac{d\hat{a}_m}{dz} = -i\beta_m \hat{a}_m - i \sum C_{mm'} \hat{a}_{m'} - \frac{\alpha_m}{2} \hat{a}_m(z) + \hat{f}_m(z). \quad (3.31)$$

where α_m is the loss term on the m th waveguide, and f_m is Langevin noise term respectively. To ensure the commutation property of the operators, the noise terms should hold the following relationship:

$$[\hat{f}_m(z), \hat{f}_{m'}^\dagger(z')] = \delta_{mm'} \alpha_m \delta(z - z'), \quad (3.32)$$

It is proved in the latter section (see 3.2.6) that the communication property is well preserved if Eq. (3.32) holds:

$$[\hat{a}_m(z), \hat{a}_{m'}^\dagger(z)] = \delta_{mm'}. \quad (3.33)$$

Similar to the two-mode case, the second order ensemble average of the noise term should fulfill [10, 19, 35]:

$$\begin{aligned}
\langle f_m^\dagger(z) f_m(z') \rangle &= \bar{n}_\sigma \delta_{mm'} \alpha_m \delta(z - z'), \\
\langle f_m(z) f_m^\dagger(z') \rangle &= (1 + \bar{n}_\sigma) \delta_{mm'} \alpha_m \delta(z - z'), \\
\bar{n}_\sigma &= \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}.
\end{aligned} \tag{3.34}$$

where ν is the beam frequency, k_B is the Boltzmann constant, and T is the temperature. At the room temperature, $\bar{n}_\sigma$ is approximately zero. The above Eq. (3.31) is a set of ordinary differential equation and the analytical solution can be obtained with the aid of Eq. (3.34):

$$\begin{aligned}
\hat{\mathbf{a}}(z) &= \mathbf{T} \hat{\mathbf{a}}(0) + \hat{\mathbf{F}}(z), \\
\mathbf{T} &= \exp(\mathbf{C}z), \\
\hat{\mathbf{F}}(z) &= \exp(\mathbf{C}z) \int_0^z \exp(-\mathbf{C}s) \hat{\mathbf{f}}(s) ds, \\
\mathbf{C} &= -\frac{\alpha}{2} - i \begin{pmatrix} \beta_1 & C_{12} & \cdots & & C_{1N} \\ C_{21} & \beta_2 & & & \\ & & \ddots & & \\ & & & \beta_{N-1} & C_{N-1N} \\ C_{N1} & & & C_{NN-1} & \beta_N \end{pmatrix}, \\
\alpha &= \begin{pmatrix} \alpha_1 & 0 & \cdots & & 0 \\ 0 & \alpha_2 & & & \\ & & \ddots & & \\ & & & \alpha_{N-1} & 0 \\ 0 & & & 0 & \alpha_N \end{pmatrix}.
\end{aligned} \tag{3.35}$$

One can calculate the correlation matrix of the Langevin noise terms in the close form:

$$\langle \hat{\mathbf{F}}(z) \hat{\mathbf{F}}^H(z) \rangle = \int_0^z \exp(\mathbf{C}s) \alpha \exp(\mathbf{C}^H s) ds, \tag{3.36}$$

where ^{H} denotes Hermitian transpose. The following property between the matrix $\mathbf{T}$ and the correlation matrix of the noise term can be derived (see 3.2.7):

$$\mathbf{T} \mathbf{T}^H + \langle \hat{\mathbf{F}}(z) \hat{\mathbf{F}}^H(z) \rangle = \mathbf{I}. \tag{3.37}$$

From a mathematical perspective, Eq. (3.37) suggests a straightforward and analytical computation for the correlation matrix of the noise terms. Physically,

Eq. (3.37) implies that, given a definite input state, the cumulative probability across all distinct eigenstates amounts to unity. It will be illustrated in the following.

The evolution of the two two-photon quantum state in waveguides n and n' can be formulated as:

$$\begin{aligned} & \hat{a}_n^\dagger(z) \hat{a}_{n'}^\dagger(z) |0\rangle \\ &= \left(\sum_{m=1}^N T_{nm}^\dagger \hat{a}_n^\dagger(0) + F_n^\dagger \right) \left(\sum_{m=1}^N T_{n'm}^\dagger \hat{a}_{n'}^\dagger(0) + F_{n'}^\dagger \right) |0\rangle, \\ & T_{nm}^\dagger = T_{nm}^*. \end{aligned} \quad (3.38)$$

If two distinguishable photons are injected, the density matrix can be computed. For example,

$$\langle 1_n^x, 1_{n'}^y | \rho | 1_n^x, 1_{n'}^y \rangle = |T_{nn} T_{n'n'}|^2. \quad (3.39)$$

where x and y denote the two polarizations of the two photons, and ρ is the density matrix.

If two indistinguishable photons are injected, that is, the polarization labeling vanishes in the equation, Eq. (3.39) should take the form $|T_{nn} T_{n'n} + T_{nn'} T_{n'n}|^2$.

One can conclude from Eqs. (3.37–3.39) that the sum of probabilities across different eigenstates, essentially the diagonal elements of the density matrix, consistently equals 1, irrespective of whether the two photons are distinguishable or indistinguishable (see 3.2.6), which suggests the theory is self-consistent.

3.3.3 Analysis of Anti-PT Symmetric Quantum Entanglement in a Three-Waveguide System

In order to demonstrate the applicability of the proposed theory, an anti-PT symmetric system three-waveguide system with one lossy waveguide is used as an example. The system can be modelled by the quantum coupled mode theory:

$$\frac{d}{dz} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} = i \begin{pmatrix} -\beta_1 & -\kappa & 0 \\ -\kappa & i\alpha - \beta_2 & -\kappa \\ 0 & -\kappa & -\beta_3 \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \end{pmatrix} + \begin{pmatrix} 0 \\ \hat{f}_2 \\ 0 \end{pmatrix}, \quad (3.40)$$

where $\beta_1, \beta_2, \beta_3$ represent the propagation constants of the respective three waveguides, while α denotes the attenuation coefficient specific to the waveguide positioned in the middle. Furthermore, κ signifies the coupling coefficient between adjacent waveguides. It is generally acceptable that the propagation constants fulfill the following:

$$\beta_1 - \beta_2 = \beta_2 - \beta_3 = \delta. \quad (3.41)$$

Under the condition that the attenuation coefficient significantly exceeds the coupling coefficient, the dynamical evolution of the annihilation operators in waveguides 1 and 3 can be elegantly described through an adiabatically simplified model [16]:

$$\begin{aligned} \frac{d}{dz} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_3 \end{pmatrix} &= \begin{pmatrix} -i\delta - \gamma - i\frac{\beta_1 + \beta_3}{2} & -\gamma \\ -\gamma & i\delta - \gamma - i\frac{\beta_1 + \beta_3}{2} \end{pmatrix} \begin{pmatrix} \hat{a}_1 \\ \hat{a}_3 \end{pmatrix}, \\ \delta &= \frac{\beta_1 - \beta_3}{2}, \\ \gamma &= \frac{|\kappa|^2}{\alpha}. \end{aligned} \quad (3.42)$$

In contrast to the simple approach embodied in Eq. (3.42), the proposed method offers a more profound understanding of the intricate physics within the waveguides. Notably, it reveals valuable insights into the photonic behavior within the middle waveguide. More crucially, it enables a comprehensive portrayal of noise terms, which can potentially be correlated after propagation.

In the follows, the parameters are specified below. The three coupled waveguides exhibit a coupling coefficient $\kappa = 0.25 \text{ cm}^{-1}$ [15], the middle waveguide experiences a loss $\alpha = 2 \text{ cm}^{-1}$, and the frequency detuning is set as $\delta = 0.8 \gamma$, indicating that the system operates in the anti-PT symmetry unbroken regime [16].

3.3.3.1 Evolution of an Indistinguishable Photon Pair

We use the quantum state $|1, 0, 1\rangle$ as an example to show how the anti-PT symmetric system impact its evolution, which can be described by the following with two indistinguishable photons:

$$|1, 0, 1\rangle = \hat{a}_1^\dagger(0)\hat{a}_3^\dagger(0)|0\rangle. \quad (3.43)$$

The operator evolves with the propagation along the waveguides:

$$\begin{aligned} \hat{a}_1^\dagger(z)\hat{a}_3^\dagger(z)|0\rangle &= \beta_{|2,0,0\rangle}|2, 0, 0\rangle + \beta_{|0,2,0\rangle}|0, 2, 0\rangle + \beta_{|0,0,2\rangle}|0, 0, 2\rangle \\ &+ \beta_{|1,0,1\rangle}|1, 0, 1\rangle + \beta_{|1,1,0\rangle}|1, 1, 0\rangle + \beta_{|0,1,1\rangle}|0, 1, 1\rangle \\ &+ \beta_{|1,0,0\rangle}|1, 0, 0\rangle + \beta_{|0,1,0\rangle}|0, 1, 0\rangle + \beta_{|0,0,1\rangle}|0, 0, 1\rangle \\ &+ \beta_{|0,0,0\rangle}|0, 0, 0\rangle, \end{aligned} \quad (3.44)$$

where:

$$\begin{aligned}
\beta_{|2,0,0\rangle} &= \sqrt{2}T_{11}^\dagger T_{31}^\dagger, \beta_{|0,2,0\rangle} = \sqrt{2}T_{12}^\dagger T_{32}^\dagger, \\
\beta_{|0,0,2\rangle} &= \sqrt{2}T_{13}^\dagger T_{33}^\dagger, \beta_{|1,0,1\rangle} = T_{11}^\dagger T_{33}^\dagger + T_{13}^\dagger T_{31}^\dagger, \\
\beta_{|1,1,0\rangle} &= T_{11}^\dagger T_{32}^\dagger + T_{12}^\dagger T_{31}^\dagger, \beta_{|0,1,1\rangle} = T_{12}^\dagger T_{33}^\dagger + T_{13}^\dagger T_{32}^\dagger, \\
\beta_{|1,0,0\rangle} &= T_{11}^\dagger F_3^\dagger + T_{31}^\dagger F_1^\dagger, \beta_{|0,1,0\rangle} = T_{12}^\dagger F_3^\dagger + T_{32}^\dagger F_1^\dagger, \\
\beta_{|0,0,1\rangle} &= T_{13}^\dagger F_3^\dagger + T_{33}^\dagger F_1^\dagger, \beta_{|0,0,0\rangle} = F_1^\dagger F_3^\dagger.
\end{aligned} \tag{3.45}$$

Obviously, the state is a mixed state, because the Langevin noise terms appears in the coefficients. Therefore, it is necessary to calculate the density matrix, which is done by Eq. (3.45). For illustration purpose, its off-diagonal element $\langle 1, 0, 0 | \rho | 0, 1, 0 \rangle$ is computed as:

$$\begin{aligned}
\langle 1, 0, 0 | \rho | 0, 1, 0 \rangle &= \left\langle \left(T_{11}^\dagger F_3^\dagger + T_{31}^\dagger F_1^\dagger \right) (T_{12} F_3 + T_{32} F_1) \right\rangle \\
&= T_{11}^\dagger T_{12} \langle F_3 F_3^\dagger \rangle + T_{31}^\dagger T_{32} \langle F_1 F_1^\dagger \rangle + T_{11}^\dagger T_{32} \langle F_1 F_3^\dagger \rangle + T_{31}^\dagger T_{12} \langle F_3 F_1^\dagger \rangle.
\end{aligned} \tag{3.46}$$

In this way, the probabilities of occupying the eigenstates, which correspond to the diagonal elements of the density matrix, can be mathematically expressed (see 3.2.8). Particularly, $p_{|1,0,1\rangle}$, the probability of detecting the indistinguishable photon pairs in waveguide 1 and waveguide 3 can be written as:

$$p_{|1,0,1\rangle} = |T_{11} T_{33} + T_{13} T_{31}|^2. \tag{3.47}$$

If the two distinguishable photons are injected, Eq. (3.47) will be $p_{|1,0,1\rangle d} = |T_{11} T_{33}|^2 + |T_{13} T_{31}|^2$.

Similar to the previous section, there is an obvious gap between the two probabilities $p_{|1,0,1\rangle}$ and $p_{|1,0,1\rangle d}$, which are plotted in Fig. 3.5a. The discrepancy is referred to as the HOM effect, with the visibility of the effect as

$$\frac{p_{|1,0,1\rangle}}{p_{|1,0,1\rangle d}} - 1. \tag{3.48}$$

The visibility of the Hong-Ou-Mandel (HOM) effect, as depicted in Fig. 3.5b, is demonstrated using both the proposed theoretical framework and the adiabatically simplified model. The close correspondence between the two curves underscores the validity and accuracy of the proposed model.

Figure 3.6 illustrates the evolutionary patterns of the probabilities within the eigenstates, specifically the terms outlined in Eq. (3.49), as a function of the propagation distance. Since the system coupling matrix is symmetric, there are overlaps for some curves (e.g., the curves of state $|1, 0, 0\rangle$ and state $|0, 0, 1\rangle$). As evident from the figure, the probability associated with state $|1, 0, 1\rangle$ diminishes as the propagation distance increases, whereas the probability of occupying the vacuum state $|0, 0, 0\rangle$

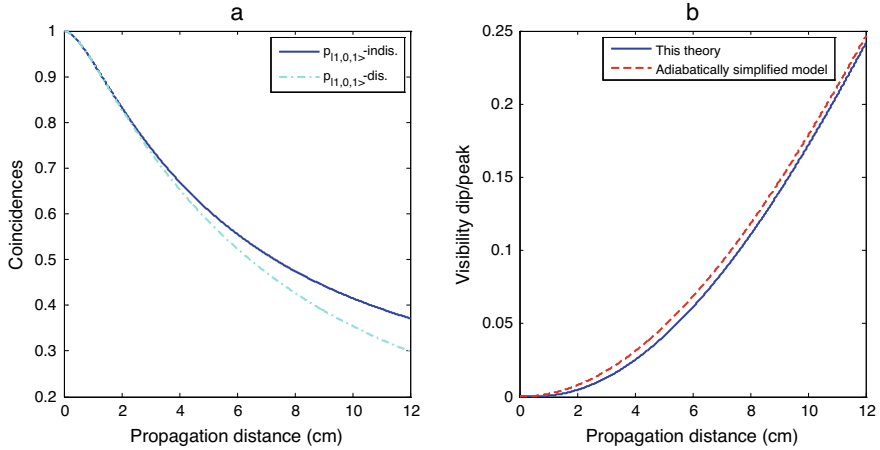


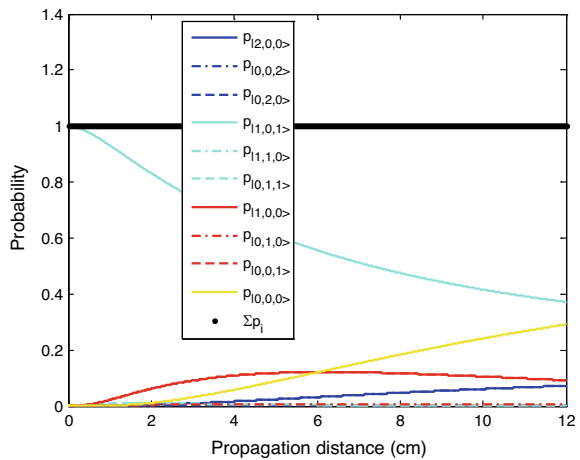
Fig. 3.5 The coincidence and the visibility for the two-photon state varying with respect to propagation distance. Dis.—stands for distinguishable photon pairs and indis.—stands for the indistinguishable photon pairs. The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]

risks correspondingly. States like $|1, 0, 0\rangle$ and $|0, 0, 1\rangle$, initially experience a probability boost, followed by a decline, which is attributed to the interplay between inter-waveguide coupling and attenuation within the middle waveguide.

As is mentioned in the previous section, the sum of the probabilities is unity, as is demonstrated in Fig. 3.6 by the black dot line:

$$p_{|2,0,0\rangle} + p_{|0,2,0\rangle} + p_{|0,0,2\rangle} + p_{|1,0,1\rangle} + p_{|0,1,1\rangle} + p_{|1,1,0\rangle} + p_{|1,0,0\rangle} + p_{|0,1,0\rangle} + p_{|0,0,1\rangle} + p_{|0,0,0\rangle} = 1. \quad (3.49)$$

Fig. 3.6 The evolution behaviors of the probabilities along the propagation distance. The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]



3.3.3.2 Evolution of an Entangled State

To further check the applicability of theory, entangled state is considered. For example, we introduce the following state:

$$|\psi(0)\rangle = \frac{1}{\sqrt{2}}(|1^x, 0, 1^y\rangle + e^{i\phi}|1^y, 0, 1^x\rangle). \quad (3.50)$$

Here, x and y represent the two distinct polarizations. The phase ϕ serves as a versatile parameter to mimic various particle behaviors [15], where $\phi = 0$ simulates bosons, and $\phi = \pi$ simulates fermions.

The entangled state will evolve in the optical waveguides:

$$|\psi(z)\rangle = \frac{1}{\sqrt{2}}(\hat{a}_1^{\dagger x}(z)\hat{a}_3^{\dagger y}(z) + e^{i\phi}\hat{a}_1^{\dagger y}(z)\hat{a}_3^{\dagger x}(z))|0\rangle, \quad (3.51)$$

which can be expanded as:

$$\begin{aligned} |\psi(z)\rangle = & \frac{1}{\sqrt{2}}(T_{11}^{\dagger}T_{31}^{\dagger} + e^{i\phi}T_{11}^{\dagger}T_{31}^{\dagger})|1^x1^y, 0, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{12}^{\dagger}T_{32}^{\dagger} + e^{i\phi}T_{12}^{\dagger}T_{32}^{\dagger})|0, 1^x1^y, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{13}^{\dagger}T_{33}^{\dagger} + e^{i\phi}T_{13}^{\dagger}T_{33}^{\dagger})|0, 0, 1^x1^y\rangle \\ & + \frac{1}{\sqrt{2}}(T_{11}^{\dagger}T_{33}^{\dagger} + e^{i\phi}T_{13}^{\dagger}T_{31}^{\dagger})|1^x, 0, 1^y\rangle \\ & + \frac{1}{\sqrt{2}}(T_{13}^{\dagger}T_{31}^{\dagger} + e^{i\phi}T_{11}^{\dagger}T_{33}^{\dagger})|1^y, 0, 1^x\rangle \\ & + \frac{1}{\sqrt{2}}(T_{12}^{\dagger}T_{33}^{\dagger} + e^{i\phi}T_{13}^{\dagger}T_{32}^{\dagger})|0, 1^x, 1^y\rangle \\ & + \frac{1}{\sqrt{2}}(T_{13}^{\dagger}T_{32}^{\dagger} + e^{i\phi}T_{12}^{\dagger}T_{33}^{\dagger})|0, 1^y, 1^x\rangle \\ & + \frac{1}{\sqrt{2}}(T_{11}^{\dagger}T_{32}^{\dagger} + e^{i\phi}T_{12}^{\dagger}T_{31}^{\dagger})|1^x, 1^y, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{12}^{\dagger}T_{31}^{\dagger} + e^{i\phi}T_{11}^{\dagger}T_{32}^{\dagger})|1^y, 1^x, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{11}^{\dagger}F_3^{\dagger y} + e^{i\phi}T_{31}^{\dagger}F_1^{\dagger y})|1^x, 0, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{31}^{\dagger}F_1^{\dagger x} + e^{i\phi}T_{11}^{\dagger}F_3^{\dagger x})|1^y, 0, 0\rangle \\ & + \frac{1}{\sqrt{2}}(T_{12}^{\dagger}F_3^{\dagger y} + e^{i\phi}T_{32}^{\dagger}F_1^{\dagger y})|0, 1^x, 0\rangle \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{\sqrt{2}} \left(T_{32}^\dagger F_1^{\dagger x} + e^{i\phi} T_{12}^\dagger F_3^{\dagger x} \right) |0, 1^y, 0\rangle \\
& + \frac{1}{\sqrt{2}} \left(T_{13}^\dagger F_3^{\dagger y} + e^{i\phi} T_{33}^\dagger F_1^{\dagger y} \right) |0, 0, 1^x\rangle \\
& + \frac{1}{\sqrt{2}} \left(T_{33}^\dagger F_1^{\dagger x} + e^{i\phi} T_{13}^\dagger F_3^{\dagger x} \right) |0, 0, 1^y\rangle \\
& + \frac{1}{\sqrt{2}} \left(F_1^{\dagger H} F_3^{\dagger y} + e^{i\phi} F_1^{\dagger V} F_3^{\dagger x} \right) |0, 0, 0\rangle.
\end{aligned} \tag{3.52}$$

In order to investigate the entangled state, the density matrix should be evaluated by considering F^y and F^x as independent variables. For illustration purpose, we can compute some elements of the density matrix as:

$$\begin{aligned}
\langle 1^x, 0, 1^y | \rho | 1^x, 0, 1^y \rangle &= \frac{1}{2} |T_{11}T_{33} + e^{-i\phi}T_{13}T_{31}|^2, \\
\langle 1^y, 0, 1^x | \rho | 1^y, 0, 1^x \rangle &= \frac{1}{2} |T_{11}T_{33} + e^{i\phi}T_{13}T_{31}|^2, \\
\langle 1^x, 0, 1^y | \rho | 1^y, 0, 1^x \rangle &= \frac{1}{2} \left(T_{11}^\dagger T_{33}^\dagger + e^{i\phi} T_{13}^\dagger T_{31}^\dagger \right) (T_{13}T_{31} + e^{-i\phi}T_{11}T_{33}), \\
\langle 0, 0, 0 | \rho | 0, 0, 0 \rangle &= \left\langle F_1 F_1^\dagger \right\rangle \left\langle F_3 F_3^\dagger \right\rangle + \text{Re} \left(\left\langle F_3 F_1^\dagger \right\rangle \left\langle F_1 F_3^\dagger \right\rangle e^{i\phi} \right).
\end{aligned} \tag{3.53}$$

One may formulate the survival probability, which is $|\langle \psi(0) | \psi(z) \rangle|$ [16, 23] as:

$$|\langle \psi(0) | \psi(z) \rangle|^2 = |T_{11}T_{33} + \cos \phi T_{13}T_{31}|^2. \tag{3.54}$$

Figure 3.7 presents a comparison between the results obtained from the proposed theory (Eq. 3.57), and the adiabatically simplified model presented in [16]. As anticipated, the survival probability undergoes a decrease with an increase in the propagation distance, attributable to the combined effects of loss and inter-waveguide coupling.

The evolutionary patterns of probabilities within distinct Eigen states have also been analyzed. Given the vast number of Eigen states, it is impractical to depict them all in Fig. 3.8. Consequently, only those states featuring two photons distributed across different waveguides are presented. System symmetry accounts for the overlap observed in some probability curves. Notably, the cumulative sum of all these probabilities, depicted by the black line, remains unity, reinforcing the coherence and consistency of the theoretical framework.

Fig. 3.7 The survival probability varying with respect to propagation distance. The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]

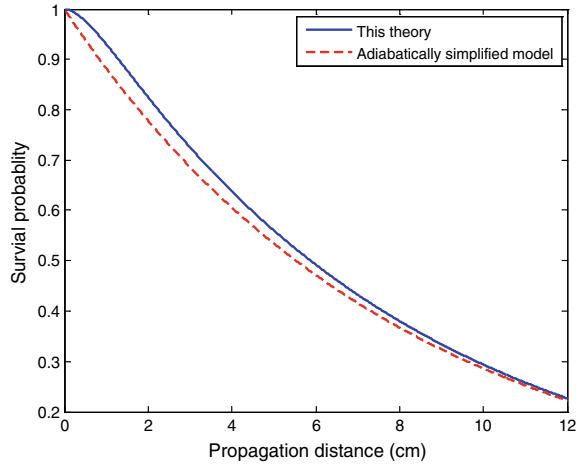
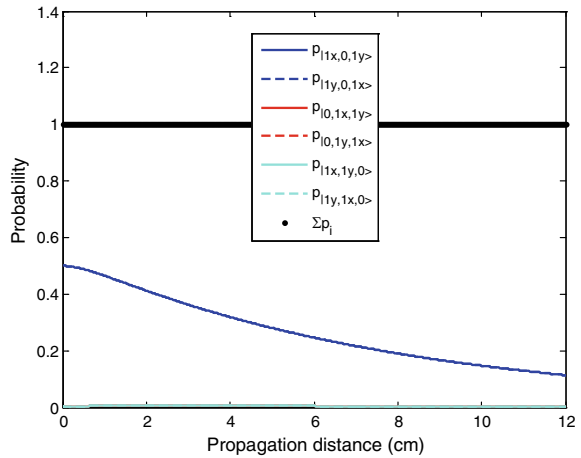


Fig. 3.8 The evolution behaviors of the probabilities along the propagation distance. The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]



3.3.4 Analysis of the Quantum State Evolution in an 89-Waveguide System

Having established the fundamental principles governing quantum state evolution in coupled lossy waveguides in the previous section, we now aim to underscore the theory's power in tackling extensive waveguide arrays. We conducted the analysis of an 89-waveguide lattice outlined in [39], which comprises 89 waveguides with a coupling coefficient of 2.9 cm^{-1} between adjacent ones. The lattice has a length of 8 mm [39], with the central waveguide designated as waveguide 0, while those positioned to its left and right are sequentially labeled from -44 to -1 and 1 to 44 , respectively [39].

The input quantum state is assumed as $\hat{a}_{-1}^\dagger \hat{a}_1^\dagger |0\rangle$ [39]. We compute the correlation Γ_{qr} [39], a metric representing the likelihood of simultaneously detecting two photons on waveguides q and r , utilizing Eqs. (3.35) and (3.39). Three distinct scenarios are explored: (a) The waveguides are devoid of any losses, (b) Each waveguide experiences an identical loss of 2 cm^{-1} , and (c) The waveguides are lossless except for the central one (waveguide 0), which has a loss of 2 cm^{-1} . These calculations are performed efficiently, taking less than 0.01 s on a PC equipped with an Intel i3-10,100 CPU running at 3.60 GHz and 8 GB of RAM. Notably, the deviation between the analytical model, defined by Eqs. (3.35) and (3.37), and the numerical model, where the 8 mm length is discretized into 5000 segments, is negligible, amounting to less than 0.03%. Furthermore, increasing the number of segments would lead to an even smaller deviation.

Figure 3.9 presents the correlation Γ_{qr} . Given that the input photons are injected into waveguides 1 and -1 , Γ_{qr} tends towards zero as the waveguide number increases. Consequently, only the correlations for waveguides ranging from -10 to 10 are depicted in Fig. 3.9. As is anticipated, the distributions of Γ_{qr} for the zero-loss and equal-loss scenarios exhibit similar patterns, with the primary difference being a reduced probability in the latter case due to losses. However, the introduction of loss specifically in the center waveguide (waveguide 0) significantly alters the pattern near this waveguide, which is observed in Fig. 3.9c as a result of the uneven absorption.

Furthermore, loss introduces a novel phenomenon that is not predicted by the lossless model: the occurrence of detecting one photon while the other is lost due to dissipation. As expected, the one-photon detection probability depicted in Fig. 3.10a is zero. In contrast, Fig. 3.10b and c illustrate the possibility of detecting one photon in the waveguides, demonstrating the impact of loss. These two figures differ not only in the probabilities but also in their distributions, attributed to the varying loss distributions employed in each scenario.

When dealing with more complicated multi-photon (>2) interactions, having the transfer matrix T at hand significantly simplifies the process of obtaining the corresponding eigenstate probabilities. For instance, in the case of a five-photon state

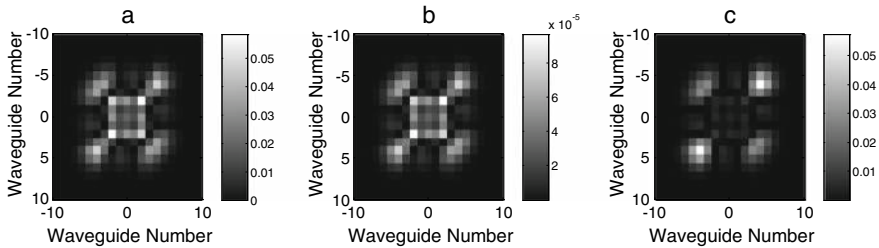


Fig. 3.9 The correlation Γ_{qr} . **a** The waveguides have no losses. **b** The waveguides have the equal loss of 2 cm^{-1} . **c** The waveguides have no losses except that the center waveguide 0 has the loss of 2 cm^{-1} . The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]

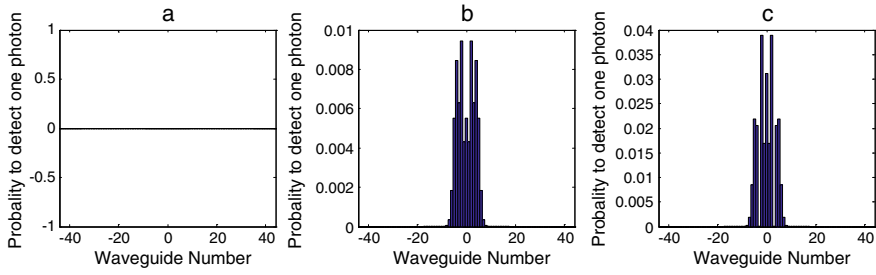


Fig. 3.10 The one photon detection probability. **a** The waveguides have no losses. **b** The waveguides have the equal loss of 2 cm^{-1} . **c** The waveguides have no losses except that the center waveguide 0 has the loss of 2 cm^{-1} . The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]

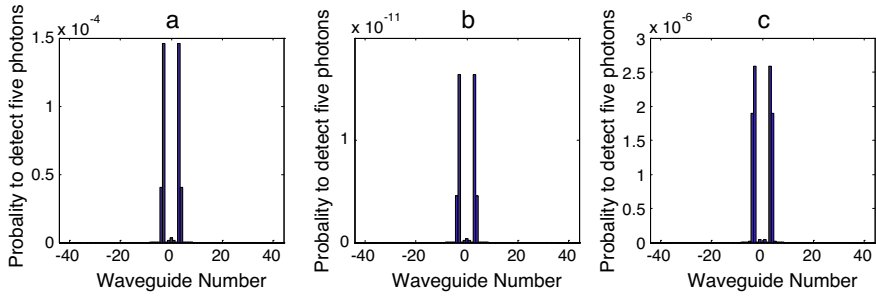


Fig. 3.11 The five-photon detection probability on each waveguide. **a** The waveguides have no losses. **b** The waveguides have the equal loss of 2 cm^{-1} . **c** The waveguides have no losses except that the center waveguide 0 has the loss of 2 cm^{-1} . The figure is reprinted/adapted with permission from [Opt. Express 30, 38,357–38,370 (2022)] © The Optical Society of America [36]

injected into waveguide 0, the probability of detecting five photons on each waveguide can be readily computed as $|T_{0m}|^0$. The outcomes of this calculation are presented in Fig. 3.11.

3.3.5 Summary

In conclusion, we have formulated a comprehensive theoretical framework for analyzing the quantum propagation of light within the coupled waveguides exhibiting unequal losses. This theory serves as a versatile tool for investigating diverse quantum phenomena, including the interplay of entangled states in the PT-symmetric and anti-PT-symmetric systems. Through straightforward matrix computations, our approach offers insightful information about the density matrix. We have applied this theory to

examine both a three-waveguide system and an 89-waveguide system, demonstrating its consistency with existing theoretical foundations.

3.3.6 Proof of the Preservation of the Commutation Property

It is mentioned that the commutation property of the annihilation and creation operators will be preserved under the proposed theoretical framework, and we will provide a detailed proof on this point.

We start with the quantum coupled mode equation:

$$\frac{d\hat{a}_m}{dz} = -i\beta_m\hat{a}_m - i \sum C_{mm'}\hat{a}_{m'} - \frac{\alpha_m}{2}\hat{a}_m(z) + \hat{f}_m(z). \quad (3.55)$$

Given the fact that the Langevin noise term commutes with the annihilation operator, we can derive:

$$d[\hat{a}_m(z), \hat{a}_m^\dagger(z)] = -\alpha_m[\hat{a}_m(z), \hat{a}_m^\dagger(z)]dz + [\hat{f}_m(z)dz, \hat{f}_m^\dagger(z)dz]. \quad (3.56)$$

For further derivation, we may evaluate the second term on the right-hand side (RHS) of Eq. (3.59) according to Eq. (3.32) as:

$$\begin{aligned} & [\hat{f}_m(z)dz, \hat{f}_m^\dagger(z)dz] \\ &= \left[\int_z^{z+dz} \hat{f}_m(s)ds, \int_z^{z+dz} \hat{f}_m^\dagger(s)ds \right] \\ &= \int_z^{z+dz} \int_z^{z+dz} [\hat{f}_m(s), \hat{f}_m^\dagger(s')]dsds' \\ &= \alpha_m dz. \end{aligned} \quad (3.57)$$

Therefore, one may have:

$$\frac{d[\hat{a}_m(z), \hat{a}_m^\dagger(z)]}{dz} = -\alpha_m[\hat{a}_m(z), \hat{a}_m^\dagger(z)]dz + \alpha_m. \quad (3.58)$$

In case m does not equal n, one has:

$$\frac{d[\hat{a}_m(z), \hat{a}_n^\dagger(z)]}{dz} = -\frac{\alpha_m}{2}[\hat{a}_m(z), \hat{a}_n(z)^\dagger] - \frac{\alpha_n}{2}[\hat{a}_m(z), \hat{a}_n(z)^\dagger]. \quad (3.59)$$

With the initial condition of:

$$[\hat{a}_m(0), \hat{a}_n^\dagger(0)] = \delta_{mn}, \quad (3.60)$$

The coupled equations (Eqs. (3.60–3.61)) can be solved analytically as:

$$[\hat{a}_m(z), \hat{a}_n^\dagger(z)] = \delta_{mn}. \quad (3.61)$$

Equation (3.61) underscores the preservation of the commutation property of the annihilation and creation operators along the propagation distance z .

3.3.7 Proof of Eq. (3.37)

In this section, we will prove the important formula in the previous section, i.e., (3.37). To start with, Eq. (3.35) is restated as:

$$\begin{aligned} \hat{\mathbf{a}}(z) &= \mathbf{T}\hat{\mathbf{a}}(0) + \hat{\mathbf{F}}(z), \\ \mathbf{T} &= \exp(\mathbf{C}z), \\ \hat{\mathbf{F}}(z) &= \exp(\mathbf{C}z) \int_0^z \exp(-\mathbf{C}s) \hat{\mathbf{f}}(s) ds. \end{aligned} \quad (3.62)$$

Relying on Eq. (3.34) and the approximation that $\bar{n}_\sigma$ is nearly zero at room temperature, one has:

$$\hat{f}(z)\hat{f}^H(z') = \alpha\delta(z - z'). \quad (3.63)$$

We may formulate the noise correlation matrix as:

$$\begin{aligned} \langle \hat{F}(z) \hat{F}^H(z) \rangle &= \exp(\mathbf{C}z) \int_0^z \exp(-\mathbf{C}s) \langle \hat{f}(s) \hat{f}^H(s') \rangle ds \int_0^z \exp(-\mathbf{C}^H s') ds' \exp(\mathbf{C}^H z) \\ &= \int_0^z \exp(\mathbf{C}(z - s)) \alpha \exp(\mathbf{C}^H(z - s)) ds \\ &= \int_0^z \exp(\mathbf{C}s) \alpha \exp(\mathbf{C}^H s) ds. \end{aligned} \quad (3.64)$$

Since the coupling matrix is symmetric, one has:

$$(\alpha = -(\mathbf{C} + \mathbf{C}^H).) \quad (3.65)$$

According to the integral's property, one has:

$$\begin{aligned}
 & \int_0^z \exp(Cs) C^H \exp(C^H s) ds \\
 &= \int_0^z \exp(Cs) d \exp(C^H s) \\
 &= \exp(Cs) \exp(C^H s) \Big|_0^z - \int_0^z (d \exp(Cs)) \exp(C^H s) \\
 &= \exp(Cz) \exp(C^H z) - I - \int_0^z \exp(Cs) C \exp(C^H s) ds. \tag{3.66}
 \end{aligned}$$

Combining Eqs. (3.64–3.66), the following can be obtained:

$$\begin{aligned}
 & TT^H + \langle \hat{F}(z) \hat{F}^H(z) \rangle \\
 &= \exp(Cz) \exp(C^H z) - \int_0^z \exp(Cs) (C + C^H) \exp(C^H s) ds \\
 &= \exp(Cz) \exp(C^H z) + I - \exp(Cz) \exp(C^H z) \\
 &= I. \tag{3.67}
 \end{aligned}$$

which is Eq. (3.37).

A direct consequence of Eq. (3.37) is that the sum of probabilities across different eigenstates, corresponding to the diagonal elements of the density matrix, always equals 1. For instance, in the case of two distinguishable photons, the sum of probabilities across these eigenstates adheres to this principle and should be:

$$\begin{aligned}
 \sum p &= \sum_{m'=1}^N \sum_{m=1}^N |T_{nm} T_{n'm'}|^2 + \sum_{m=1}^N |T_{nm}|^2 \langle F_{n'} F_{n'}^\dagger \rangle \\
 &+ \sum_{m=1}^N |T_{nm}|^2 \langle F_n F_n^\dagger \rangle + \langle F_n F_n^\dagger \rangle \langle F_{n'} F_{n'}^\dagger \rangle \\
 &= \left(\sum_{m'=1}^N |T_{n'm'}|^2 + \langle F_{n'} F_{n'}^\dagger \rangle \right) \left(\sum_{m=1}^N |T_{nm}|^2 + \langle F_n F_n^\dagger \rangle \right) \\
 &= 1 \times 1 \\
 &= 1. \tag{3.68}
 \end{aligned}$$

The same principle should take effect in case the two photons are indistinguishable:

$$\begin{aligned}
& \left\langle \frac{1}{2} \sum_{m'=1}^N \sum_{m=1}^N \left| T_{nm}^\dagger T_{n'm'}^\dagger + T_{nm'}^\dagger T_{n'm}^\dagger \right|^2 \right\rangle + \left\langle \sum_{m=1}^N \left| T_{nm}^\dagger F_{n'}^\dagger + T_{n'm}^\dagger F_n^\dagger \right|^2 \right\rangle \\
& + \left\langle F_{n'}^\dagger F_n \right\rangle \left\langle F_{n'} F_n^\dagger \right\rangle + \left\langle F_n^\dagger F_n \right\rangle \left\langle F_{n'} F_{n'}^\dagger \right\rangle \\
& = \left| \sum_{m=1}^N T_{n'm}^\dagger T_{nm} + \left\langle F_{n'}^\dagger F_n \right\rangle \right|^2 + \left(\sum_{m=1}^N |T_{n'm}|^2 + \left\langle F_{n'} F_{n'}^\dagger \right\rangle \right) \\
& \quad \left(\sum_{m=1}^N |T_{nm}|^2 + \left\langle F_n F_n^\dagger \right\rangle \right) \\
& = 0 + 1 \\
& = 1.
\end{aligned} \tag{3.69}$$

3.3.8 Derivation of the Diagonal Terms of the Density Matrix with Input as $|1, 0, 1\rangle$

In this section, we consider the propagation of state $|1, 0, 1\rangle$ in the three coupled waveguides. The probabilities associated with the various eigenstates, represented by the diagonal terms of the density matrix, can be formulated as:

$$\begin{aligned}
p_{|2,0,0\rangle} &= 2|T_{11}T_{31}|^2, \\
p_{|0,2,0\rangle} &= 2|T_{12}T_{32}|^2, \\
p_{|0,0,2\rangle} &= 2|T_{13}T_{33}|^2, \\
p_{|1,0,1\rangle} &= |T_{11}T_{33} + T_{13}T_{31}|^2, \\
p_{|1,1,0\rangle} &= |T_{11}T_{32} + T_{12}T_{31}|^2, \\
p_{|0,1,1\rangle} &= |T_{12}T_{33} + T_{13}T_{32}|^2, \\
p_{|1,0,0\rangle} &= |T_{11}F_3 + F_1T_{31}|^2 = |T_{11}|^2 \left\langle F_3 F_3^\dagger \right\rangle \\
& \quad + |T_{31}|^2 \left\langle F_1 F_1^\dagger \right\rangle + 2\text{Re} \left(T_{11}T_{31}^\dagger \left\langle F_3 F_1^\dagger \right\rangle \right), \\
p_{|0,1,0\rangle} &= |T_{12}F_3 + F_1T_{32}|^2 = |T_{12}|^2 \left\langle F_3 F_3^\dagger \right\rangle \\
& \quad + |T_{32}|^2 \left\langle F_1 F_1^\dagger \right\rangle + 2\text{Re} \left(T_{12}T_{32}^\dagger \left\langle F_3 F_1^\dagger \right\rangle \right), \\
p_{|0,0,1\rangle} &= |T_{13}F_3 + F_1T_{33}|^2 = |T_{13}|^2 \left\langle F_3 F_3^\dagger \right\rangle
\end{aligned}$$

$$\begin{aligned}
& + |T_{33}|^2 \langle F_1 F_1^\dagger \rangle + 2\text{Re} \left(T_{13} T_{33}^\dagger \langle F_3 F_1^\dagger \rangle \right), \\
p_{|0,0,0\rangle} &= \langle F_1 F_1^\dagger \rangle \langle F_3 F_3^\dagger \rangle + \langle F_3 F_1^\dagger \rangle \langle F_1 F_3^\dagger \rangle.
\end{aligned} \tag{3.70}$$

where Re denotes the real part of the term. If the two photons are distinguishable, the probabilities can be computed as:

$$\begin{aligned}
p_{|2,0,0\rangle_d} &= |T_{11} T_{31}|^2, \\
p_{|0,2,0\rangle_d} &= |T_{12} T_{32}|^2, \\
p_{|0,0,2\rangle_d} &= |T_{13} T_{33}|^2, \\
p_{|1,0,1\rangle_d} &= |T_{11} T_{33}|^2 + |T_{13} T_{31}|^2, \\
p_{|1,1,0\rangle_d} &= |T_{11} T_{32}|^2 + |T_{12} T_{31}|^2, \\
p_{|0,1,1\rangle_d} &= |T_{12} T_{33}|^2 + |T_{13} T_{32}|^2, \\
p_{|1,0,0\rangle_d} &= |T_{11} F_3 + F_1 T_{31}|^2 = |T_{11}|^2 \langle F_3 F_3^\dagger \rangle + |T_{31}|^2 \langle F_1 F_1^\dagger \rangle, \\
p_{|0,1,0\rangle_d} &= |T_{12} F_3 + F_1 T_{32}|^2 = |T_{12}|^2 \langle F_3 F_3^\dagger \rangle + |T_{32}|^2 \langle F_1 F_1^\dagger \rangle, \\
p_{|0,0,1\rangle_d} &= |T_{13} F_3 + F_1 T_{33}|^2 = |T_{13}|^2 \langle F_3 F_3^\dagger \rangle + |T_{33}|^2 \langle F_1 F_1^\dagger \rangle, \\
p_{|0,0,0\rangle_d} &= \langle F_1 F_1^\dagger \rangle \langle F_3 F_3^\dagger \rangle.
\end{aligned} \tag{3.71}$$

References

1. J.H. Zhou, P. Gallion, Comprehensive analytical model to characterize randomness in optical waveguides. *Opt. Express* **24**(7), 6825–6842 (2016)
2. A. Snyder, J. Love, *Optical Waveguide Theory* (Springer, 1983)
3. D. Marcuse, Mode conversion caused by surface imperfections of a dielectric slab waveguide. *Bell Syst. Tech. J.* **48**(10), 3187–3215 (1969)
4. D. Marcuse, Derivation of coupled power equations. *Bell Syst. Tech. J.* **51**(1), 229–237 (1972)
5. D. Marcuse, Power distribution and radiation losses in multimode dielectric slab waveguides. *Bell Syst. Tech. J.* **51**(2), 429–454 (1972)
6. M. Huang, X. Yan, Thermal-stress effects on the temperature sensitivity of optical waveguides. *J. Opt. Soc. Am. B* **20**(6), 1326–1333 (2003)
7. J. Seligson, The orthogonality relation for TE- and TM-modes in guided-wave optics. *J. Lightwave Technol.* **6**(8), 1260–1264 (1988)
8. W.-P. Huang, Coupled-mode theory for optical waveguides: an overview. *J. Opt. Soc. Am. A* **11**(3), 963–983 (1994)
9. D. Marcuse, Coupled-mode theory for anisotropic optical waveguides. *Bell Syst. Tech. J.* **54**(6), 985–995 (1975)
10. J.H. Zhou, Characterization of PT-symmetric quantum interference based on the coupled mode theory. *Opt. Express* **30**(13), 23600–23607 (2022)
11. G. Moody et al., Roadmap on integrated quantum photonics. *J. Phys.: Photon.* **4** (2022)
12. E. Saglamyurek et al., Broadband waveguide quantum memory for entangled photons. *Nature* **469**(7331), 512–515 (2011)

13. A. Javadi et al., Single-photon non-linear optics with a quantum dot in a waveguide. *Nat. Commun.* **6**(1), 8655 (2015)
14. T. Kobayashi et al., Frequency-domain Hong–Ou–Mandel interference. *Nat. Photonics* **10**(7), 441–444 (2016)
15. F. Klauck, L. Teuber, M. Ornigotti, M. Heinrich, S. Scheel, A. Szameit, Observation of PT-symmetric quantum interference. *Nat. Photonics* **13**(12), 883–887 (2019)
16. Y. Qin et al., Quantum interference in anti-parity-time symmetric coupled waveguide system. *Opt. Express* **29**(18), 29175–29185 (2021)
17. M. Ehrhardt, M. Heinrich, A. Szameit, Observation-dependent suppression and enhancement of two-photon coincidences by tailored losses. *Nat. Photonics* **16**(3), 191–195 (2022)
18. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic Press, New York, 1974)
19. M.O. Scully, M.S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, 1997)
20. A. Lukš, V. Peřinová, *Canonical Quantum Description of Light Propagation in Dielectric Media (Progress in Optics)* (Elsevier, New York, 2002), pp.295–431
21. J. Liñares, M.C. Nistal, Quantization of coupled modes propagation in integrated optical waveguides. *J. Mod. Opt.* **50**, 781–790 (2003)
22. J. Liñares, M.C. Nistal, D. Barral, Quantization of coupled 1D vector modes in integrated photonic waveguides. *New J. Phys.* **10**, 063023 (2008)
23. D.B. Raña, *Spatial Propagation and Characterization of Quantum States of Light in Integrated Photonic Devices*, Universidade De Santiago De Compostela (2015)
24. A. Rai, S. Das, G.S. Agarwal, Quantum entanglement in coupled lossy waveguides. *Opt. Express* **18**(6), 6241–6254 (2010)
25. C.M. Bender, S. Boettcher, Real spectra in non-Hermitian Hamiltonians having PT symmetry. *Phys. Rev. Lett.* **80**(24), 5243–5246 (1998)
26. Ş.K. Özdemir, S. Rotter, F. Nori, L. Yang, Parity–time symmetry and exceptional points in photonics. *Nat. Mater.* **18**(8), 783–798 (2019)
27. R. El-Ganainy, K.G. Makris, D.N. Christodoulides, Z.H. Musslimani, Theory of coupled optical PT-symmetric structures. *Opt. Lett.* **32**(17), 2632–2634 (2007)
28. C.E. Rüter, K.G. Makris, R. El-Ganainy, D.N. Christodoulides, M. Segev, D. Kip, Observation of parity–time symmetry in optics. *Nat. Phys.* **6**(3), 192–195 (2010)
29. Y. Li et al., Coupled mode theory under the parity-time symmetry frame. *J. Lightwave Technol.* **31**, 2477–2481 (2013)
30. A. Cerjan, S. Fan, Eigenvalue dynamics in the presence of nonuniform gain and loss. *Phys. Rev. A* **94**(3), 033857 (2016)
31. Z.-Z. Liu, Q. Zhang, Y. Chen, J.-J. Xiao, General coupled-mode analysis of a geometrically symmetric waveguide array with nonuniform gain and loss. *Photon. Res.* **5**(2), 57–63 (2017)
32. X. Zhang, V.A. Vysloukh, Y.V. Kartashov, X. Chen, F. Ye, M.R. Belić, PT symmetry in nonlinear twisted multicore fibers. *Opt. Lett.* **42**(15), 2972–2975 (2017)
33. C.K. Hong, Z.Y. Ou, L. Mandel, Measurement of sub-picosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**(18), 2044–2046 (1987)
34. B. Vest et al., Anti-coalescence of bosons on a lossy beam splitter. *Science* **356**(6345), 1373–1376 (2017)
35. S. Jiang, *Contributions à l'étude théorique des bruits quantiques et classiques dans les amplificateurs Raman distribués*, Telecom ParisTech (2008)
36. J.H. Zhou, Analytical formulation of quantum interference inside coupled waveguides with unequal losses. *Opt. Express* **30**(21), 38357–38370 (2022)
37. F. Yang, Y.-C. Liu, L. You, Anti-PT symmetry in dissipatively coupled optical systems. *Phys. Rev. A* **96**(5), 053845 (2017)
38. S. Weimann et al., Implementation of quantum and classical discrete fractional Fourier transforms. *Nat. Commun.* **7**(1), 11027 (2016)
39. Y. Bromberg, Y. Lahini, R. Morandotti, Y. Silberberg, Quantum and classical correlations in waveguide lattices. *Phys. Rev. Lett.* **102**(25), 253904 (2009)
40. S.M. Barnett, J. Jeffers, A. Gatti, R. Loudon, Quantum optics of lossy beam splitters. *Phys. Rev. A* **57**(3), 2134–2145 (1998)

41. D.N. Makarov, Theory of a frequency-dependent beam splitter in the form of coupled waveguides. *Sci. Rep.* **11**(1), 5014 (2021)
42. D.N. Makarov, Theory of HOM interference on coupled waveguides. *Opt. Lett.* **45**(22), 6322–6325 (2020)

Chapter 4

Wave Propagation in Multi-core Fibers



Abstract Multi-core fibers. The chapter studies the multi-core fibers, which are considered as a special kind of multi-mode waveguide while the cores are closely coupled to each other and form the super-modes. The concept is discussed and analytical solutions are given for various multi-core fibers with different core arrangement.

Multi-core fibers (MCFs) have recently emerged as a new candidate to further increase the transmission capacity of the optical communication systems [1–8]. In the MCF based optical transmission systems, one implements the modes of individual cores or the super-modes of the MCF to transmit optical signals in order to achieve a higher spectral efficiency in comparison with the single core fiber transmission systems [1, 2, 4, 9–14]. It is reported that a MCF with 19 cores have been successfully fabricated [5] and experimentally demonstrated during the transmission [6]. Apart from the optical communication systems, MCFs are also useful for fiber lasers [15–18], passive optical switch and star couplers [19], and fiber endoscope [20].

As for wave propagation inside the multi-core fibers, consideration should be taken on the mode coupling effect between the cores, which makes multi-core fibers a special kind of multi-mode waveguides. Particularly, while the cores are closely coupled to each other, it will form the super-modes. The concept is discussed and analytical solutions are given for various multi-core fibers with different core arrangement.

4.1 Analytical Formulation of Super-Modes Inside Multi-core Fibers with Circularly Distributed Cores

Part of the section reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21].

This section delves into a comprehensive analysis of super-modes within multi-core fibers, featuring cores distributed in circular patterns. These patterns encompass arrangements including a single ring, dual rings, and multiple rings of cores. Additionally, the discussion encompasses multi-core fibers (MCFs) that incorporate a central core embedded amidst these ring structures. Within this framework, analytical expressions are formulated for both the propagation constants and the modal distribution vectors of the super-modes, providing a quantitative understanding of their characteristics.

4.1.1 Introduction

To effectively integrate multi-core fibers (MCFs) into various applications, a thorough examination of the super-mode properties is essential. Specifically, in a mode-division multiplexing (MDM) system utilizing MCFs, the information is transmitted through the super-modes rather than relying solely on individual core modes. This approach can notably increase the spectral efficiency within optical fiber communication systems [2, 12, 15]. Furthermore, in MCF-based fiber lasers, the near-field and far-field characteristics, as well as the longitudinal mode wavelengths, are intimately tied to the modal distributions and effective indices of the super-modes [15, 16]. This underscores the criticality of thoroughly understanding and manipulating these super-mode properties for optimal performance.

Recently, there has been a surge of publications dedicated to exploring the super-modes of multi-core fibers (MCFs) and fiber arrays [2, 12, 15–17, 19, 22–24]. A portion of these studies employed numerical simulations [2, 12, 15] to illustrate the modal behavior within MCFs to investigate the characteristics of their super-modes. In parallel, others have analytically studied the eigenmodes and propagation constants of super-modes in MCFs and fiber arrays [17, 19, 22, 23]. Compared to numerical findings, analytical expressions provide a profound understanding of the fundamental physics governing super-modes. They offer a more concise and intuitive picture for the design and analysis of MCFs, making them highly appealing to both academic researchers and industrial practitioners. As such, these analytical approaches have gained significant interest across both sectors.

Currently, analytical expressions for propagation constants and eigenmode vectors have been established for MCFs with linearly aligned cores [17, 19] and circularly arranged cores within a single ring [19, 22]. However, to optimize core alignment and maximize the benefits of MCFs, distributing cores across multiple rings (e.g., two [23], three, or more) has become a preferred approach. This arrangement increases the number of cores within a single MCF, thereby boosting the number of super-modes and transmission capacity in MDM systems or increasing the output power in fiber lasers while maintaining high beam quality. Moreover, incorporating an additional core at the center of the rings in MCFs offers remarkable benefits. In the MCF-based MDM systems, the central core can enhance the overall spectral efficiency of the fiber [2]; In the MCF-based lasing systems, the presence of the center core enlarges

the mode area, resulting in a substantial boost in laser output power [15]; for MCF based star couplers [19], the central core plays a pivotal role in improving routing directivity, ensuring efficient signal distribution. Despite these advantages, there are few analytical discussions exploring the behavior of MCFs with cores arranged in multiple rings, including those with a central core. Ref. [23] presented an analytical mode for the two-ring MCFs by simply assuming that the coupling coefficients between adjacent cores within both rings are uniform, which is less accurate in case the rings have large diameters. A one-ring MCF with a center core was investigated analytically in [24], but the formulas have not been generalized for the multiple ring case.

In this section, we conduct the analysis of the super-modes inside the multi-core fibers (MCFs) with the cores arranged in multiple concentric rings. The study focusses both on the propagation constants and the eigenmode distributions, which we derive precise analytical expressions. To validate the accuracy of our analytical findings, we conduct numerical simulations based on the beam propagation method (BPM), thereby ensuring the credibility of our results. This comprehensive approach ensures a deep understanding of the behavior of MCFs with complex core arrangements.

In the subsequent analysis, it is important to note that we assume coupling occurs exclusively between adjacent fiber cores, and furthermore, that the number of cores within each individual ring remains constant.

4.1.2 MCFs with Cores in One Ring

Janice Hudgings et al. [19] has analyzed the fundamental case, i.e., the MCFs with a one-ring fiber core arrangement. Given that all cores are assumed to be uniform and single-mode, the field distribution within each core should take the form $a_k \exp(-j\beta z)$ [19], where a_k represents the amplitude of the field in the k th core and β denotes the common propagation constant shared by all the single-mode cores. In accordance with Ref. [2, 19], the amplitudes of the individual cores must satisfy the following coupled mode equation.

$$\begin{aligned}
\frac{d\mathbf{A}}{dz} &= -j\kappa\mathbf{A} \\
\mathbf{A} &= \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix} \\
\kappa &= \begin{pmatrix} 0 & \kappa & & \kappa \\ \kappa & 0 & \kappa & \\ & \ddots & \ddots & \ddots \\ & & \kappa & 0 & \kappa \\ \kappa & & & \kappa & 0 \end{pmatrix}
\end{aligned} \tag{4.1}$$

where N represents the total number of cores within the ring, κ denotes the coupling coefficients between the adjacent cores [2, 19].

It can be easily inferred from Eq. (4.1) that matrix κ is symmetric and Hermitian, which enables it to be decomposed into $\mathbf{Q}\mathbf{D}\mathbf{Q}^H$, where $\mathbf{Q}$ is orthogonal or unitary depending on whether κ is real or complex, $\mathbf{D}$ is diagonal with its diagonal terms as the Eigen value of κ , and H denotes the Hermitian operation. Based on the above analysis, we have the solution of the coupled mode equation as [25]

$$\mathbf{A}(L) = \mathbf{Q} \exp(-j\mathbf{D}L) \mathbf{Q}^H \mathbf{A}(0) \tag{4.2}$$

It is evident that the columns of the matrix $\mathbf{Q}$ constitute the eigenvectors, representing the super-mode of the system. Correspondingly, the values $\beta + d_n$, where d_n is the n th diagonal element of the matrix $\mathbf{D}$, function as the propagation constants for the respective super-modes.

It was already derived analytically in [19] that the propagation constants of the super-modes is as follows:

$$\beta + 2\kappa \cos\left(\frac{2\pi(n-1)}{N}\right) \tag{4.3}$$

where n represents the order or index of the super-modes and one may analytically formulate the element of matrix $\mathbf{Q}$ as [19]

$$Q_{mn} = \sqrt{\frac{1}{N}} \exp\left(-j(m-1)(n-1)\frac{2\pi}{N}\right) \tag{4.4}$$

Given that Eq. (4.4) explicitly presents the eigenvectors of the system in a complex form, it is essential to recognize that in practical numerical simulations, the mode amplitudes are commonly represented by real values. To facilitate the analysis, the matrix $\mathbf{Q}$ is transformed from the complex one into the real one:

when N is an even number

$$Q_{mn} = \begin{cases} \sqrt{\frac{1}{N}} & n = 1 \\ \sqrt{\frac{2}{N}} \cos((m-1)(n-1)\frac{2\pi}{N}) & 1 < n \leq \frac{N}{2} \\ \sqrt{\frac{1}{N}}(-1)^{m-1} & n = 1 + \frac{N}{2} \\ \sqrt{\frac{2}{N}} \sin((m-1)(N-n+1)\frac{2\pi}{N}) & \frac{N}{2} + 1 < n \leq N \end{cases} \quad (4.5)$$

when N is an odd number

$$Q_{mn} = \begin{cases} \sqrt{\frac{1}{N}} & n = 1 \\ \sqrt{\frac{2}{N}} \cos((m-1)(n-1)\frac{2\pi}{N}) & 1 < n \leq \frac{N+1}{2} \\ \sqrt{\frac{2}{N}} \sin((m-1)(N-n+1)\frac{2\pi}{N}) & \frac{N+1}{2} < n \leq N \end{cases}$$

If $\mathbf{Q}$ is real, the matrix property ensures $\mathbf{Q}^H = \mathbf{Q}^T$, where T denotes transpose operation on the matrix. The validity of Eq. (4.5) can be proved by comparing with the results in [2], which demonstrates a precise matching. Therefore, we may infer that $\mathbf{Q}$ is not unique. This can also be observed from the expression of real $\mathbf{Q}$ that mode n and mode $N + 2 - n$ are degenerated with identical propagation constant. This suggests that these modes can be formulated by the super-position of them degenerated modes.

It is important to emphasize that the matrix $\mathbf{Q}$ is independent of the coupling coefficient κ , and it can be applied to the MCFs with one ring uniform cores albeit with different parameters.

4.1.3 MCFs with Cores in Two Rings

4.1.3.1 Mathematical Formulation

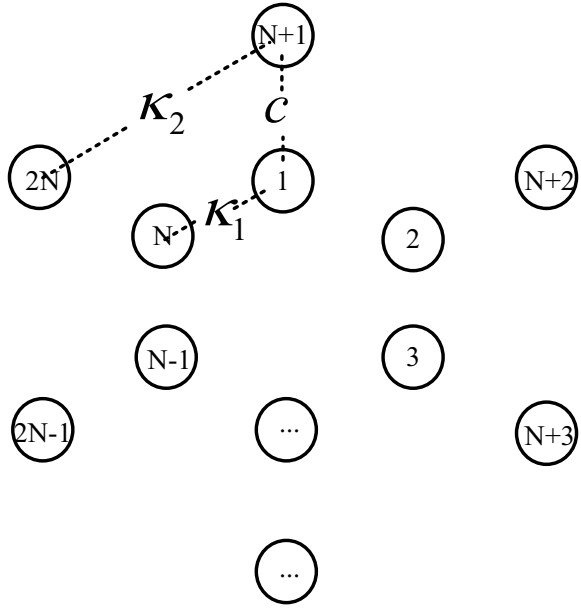
In this analysis, we maintain the assumption that mode coupling within each individual ring of cores is uniform, although the coupling coefficients between cores belonging to the two distinct rings, i. e. κ_1 and κ_2 , are not identical. Furthermore, we acknowledge the existence of coupling between adjacent cores situated in different rings, and this inter-ring coupling is characterized by a coefficient c , which is not identical to κ_1 or κ_2 .

The arrangement of the cores within the MCFs is shown in Fig. 4.1.

Under these assumptions, we may formulate the coupled mode equation as

$$\begin{aligned} \frac{d\mathbf{A}}{dz} &= -j\mathbf{M}\mathbf{A} \\ \mathbf{M} &= \begin{pmatrix} \kappa_1 & \mathbf{C} \\ \mathbf{C} & \kappa_2 \end{pmatrix} \end{aligned} \quad (4.6)$$

Fig. 4.1 The arrangement of the fiber array and the coupling between the cores. This figure is reprinted/ adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]



where the matrices κ_1 and κ_2 represent the coupling interactions between cores within their respective rings, the diagonal matrix C is composed of the coupling coefficients between cores belonging to different rings. Since the Eigen vectors of κ_1 and κ_2 are independent of coupling coefficients κ_1 and κ_2 , they can be formulated by singular value decomposition as

$$\begin{aligned}\kappa_1 &= \mathbf{Q}\mathbf{D}_1\mathbf{Q}^H \\ \kappa_2 &= \mathbf{Q}\mathbf{D}_2\mathbf{Q}^H\end{aligned}\tag{4.7}$$

Therefore, the matrix $\mathbf{M}$ can be formulated in the block matrix sense as:

$$\begin{aligned}\mathbf{M} &= \begin{pmatrix} \mathbf{Q}\mathbf{D}_1\mathbf{Q}^H & \mathbf{Q}\mathbf{C}\mathbf{Q}^H \\ \mathbf{Q}\mathbf{C}\mathbf{Q}^H & \mathbf{Q}\mathbf{D}_2\mathbf{Q}^H \end{pmatrix} \\ &= \mathbf{Q}_{Total} \begin{pmatrix} \mathbf{D}_1 & \mathbf{C} \\ \mathbf{C} & \mathbf{D}_2 \end{pmatrix} \mathbf{Q}_{Total}^H \\ \mathbf{Q}_{Total} &= \begin{pmatrix} \mathbf{Q} & 0 \\ 0 & \mathbf{Q} \end{pmatrix}\end{aligned}\tag{4.8}$$

To further derive the expression for the modes and the propagation constants, further matrix decomposition is employed on the matrix in the middle of the expression:

$$\begin{pmatrix} \mathbf{D}_1 & \mathbf{C} \\ \mathbf{C} & \mathbf{D}_2 \end{pmatrix} = \mathbf{U}\mathbf{N}\mathbf{U}^H \quad (4.9)$$

where

$$\mathbf{N} = \begin{pmatrix} d_1^1 & c & & \\ c & d_2^1 & & \\ & & \ddots & \\ & & & d_1^N & c \\ & & & c & d_2^N \end{pmatrix} \quad (4.10)$$

In this scenario, the matrix $\mathbf{U}$ is constructed by modifying the identity matrix through a specific pattern of row exchanges. The k th row of the identity matrix is exchanged with the $(2k-1)$ th row when k is between $[1, N]$ or is exchanged with the $2(k-N)$ th row when k is between $[N+1, 2N]$.

Through singular value decomposition, the symmetric matrix $\mathbf{N}$ can be formulated as:

$$\mathbf{N} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \quad (4.11)$$

The diagonal matrix $\mathbf{\Lambda}$ is constructed so that its elements are related to the eigenvalues of two matrices, i.e., $\mathbf{M}$ and $\mathbf{N}$. Particularly its $(2n-1)$ th element and $(2n)$ th element are the Eigen values of the sub-matrix $\begin{pmatrix} d_1^n & c \\ c & d_2^n \end{pmatrix}$. The Eigen values can be computed analytically by the following equation

$$\lambda^2 - (d_1^n + d_2^n)\lambda + d_1^n d_2^n - c^2 = 0 \quad (4.12)$$

and the roots of Eq. (4.12) are evaluated as

$$\lambda = \frac{(d_1^n + d_2^n) \pm \sqrt{(d_1^n - d_2^n)^2 + 4c^2}}{2} \quad (4.13)$$

Therefore, the propagation constants of the super-modes should be

$$\beta + (\kappa_1 + \kappa_2) \cos\left(\frac{2\pi(n-1)}{N}\right) \pm \sqrt{(\kappa_1 - \kappa_2)^2 \cos^2\left(\frac{2\pi(n-1)}{N}\right) + c^2} \quad (4.14)$$

The related eigenvectors for the sub-matrix are.

$$\begin{aligned}
& \sqrt{\frac{c^2}{2k^2 + 2c^2 \pm 2k\sqrt{k^2 + c^2}}} \\
& \times \left(\frac{k \pm \sqrt{k^2 + c^2}}{c}, 1 \right) \\
& k = (\kappa_1 - \kappa_2) \cos\left(\frac{2\pi(n-1)}{N}\right)
\end{aligned} \tag{4.15}$$

Therefore, the matrix $\mathbf{V}$ will be composed by the block matrixes whose elements are formed by the two orthogonal vectors in Eq. (4.15). Totally, the Eigen vectors of matrix $\mathbf{M}$, which describes the modal amplitudes distribution within the cores, will be the rows of the matrix $\mathbf{Q}_{\text{total}}\mathbf{UV}$. The $(2n-1)$ th and the $(2n)$ th Eigen vectors, i. e. the corresponding rows of the matrix $\mathbf{Q}_{\text{total}}\mathbf{UV}$, will be the row vectors of the following $N \times 2$ matrix

$$\begin{pmatrix} \mathbf{q}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{q}_n \end{pmatrix} \begin{pmatrix} v_n^{11} & v_n^{12} \\ v_n^{21} & v_n^{22} \end{pmatrix} = \begin{pmatrix} v_n^{11} \mathbf{q}_n & v_n^{12} \mathbf{q}_n \\ v_n^{21} \mathbf{q}_n & v_n^{22} \mathbf{q}_n \end{pmatrix} \tag{4.16}$$

where v_{ij} is the element on the i th column and j th row of $\mathbf{V}$. The expression suggests that the modes of the two-ring system is actually the superposition of the modes within the individual rings.

4.1.3.2 Two Special Cases

First of all, we assume that the coupling coefficient among the different rings $c = 0$. Therefore, we have Eq. (4.14) as:

$$\beta + (\kappa_1 + \kappa_2) \cos\left(\frac{2\pi(n-1)}{N}\right) \pm (\kappa_1 - \kappa_2) \cos\left(\frac{2\pi(n-1)}{N}\right) \tag{4.17}$$

The results obtained under the condition where $c = 0$ coincide exactly with those predicted by Eq. (4.3). The reason for this is that when $c = 0$, the two rings operate separately and Eq. (4.3) for the individual ring should apply.

When the three coupling coefficients κ_1 , κ_2 and c are identical, Eq. (4.14) becomes

$$\beta + \left(2\kappa_1 \cos\left(\frac{2\pi(n-1)}{N}\right) \right) \pm \kappa_1 \tag{4.18}$$

The results are in line with the findings reported in [17], which adopted the assumption of equal coupling coefficients between adjacent cores, both within the same ring and between different rings.

In this case, the Eigen vectors in Eq. (4.15) should take the form

$$\left(\pm \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)^T \quad (4.19)$$

Therefore, the corresponding Eigen modes of the entire multicore fiber (MCF) system should be in the form of the following:

$$\left(\pm \frac{1}{\sqrt{2}} \mathbf{q}^T, \frac{1}{\sqrt{2}} \mathbf{q}^T \right)^T \quad (4.20)$$

This is in precise alignment with the results presented in [17].

4.1.3.3 Numerical Verifications

The three-dimensional Beam Propagation Method (3DBPM) has been employed to validate the proposed theoretical framework for a specific six-core, two-ring multi-core fiber (MCF) configuration. As depicted in Fig. 4.2, the six cores are arranged in a pattern consisting of two triangles, with the cores within the rings separated by the distances $d_1 = 6.0622 \mu\text{m}$ and $d_2 = 17.3205 \mu\text{m}$ respectively. Additionally, the adjacent cores belonging to different rings are separated by $d = 6.5 \mu\text{m}$. Apart from the unique core arrangement, the physical properties of the cores and cladding, such as the refractive indexes and the sizes of the cores, are consistent with those reported [10]. All cores possess equal diameters of $a = 5 \mu\text{m}$ [10]. The refractive index difference between the cores and cladding is 1.2%, with the cladding index being 1.45 [10]. The operation wavelength of the MCF is 1550 nm.

To validate the proposed analytical formulas, the effective indices of the super-modes are calculated utilizing both Eq. (4.14) and the Beam Propagation Method (BPM). The computation of the effective indexes is done by dividing the propagation constants of the respective modes by the free space wave number k_0 . The

Fig. 4.2 The 6-core-MCF core array arrangement. This figure is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

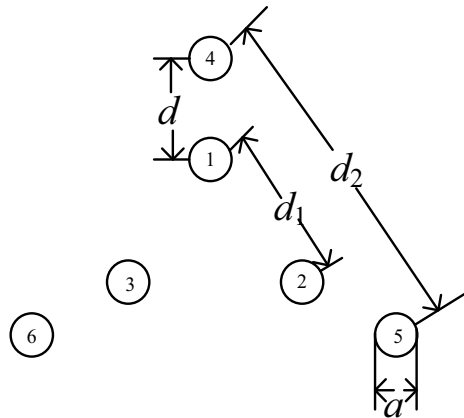


Table 4.1 The effective indexes of the supermodes obtained by the both methods

Super-mode number	BPM results	Results from Eq. (14)
1	1.4596	1.4598
2	1.4572	1.4572
3	1.4572	1.4572
4	1.4563	1.4562
5	1.4545	1.4546
6	1.4545	1.4546

This table is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

coupling coefficients between the cores, which are essential inputs to Eq. (4.14), are numerically acquired based on the coupled mode theory [26].

The results comparison is presented in Table 4.1, where it is evident that the outcomes derived from Eq. (4.14) align exceptionally well with the predictions made by the Beam Propagation Method (BPM). The discrepancy between these two sets of results is minimal, with a margin of error below 0.15%, demonstrating the strong agreement between the theoretical calculations and the numerical simulations.

Subsequent to calculating the propagation constants of the super-modes, we conduct the analysis of the amplitude distribution across each of the cores for these super-modes.

We depict the BPM simulations of the super-mode field in Fig. 4.3. The results reveal that the super-modes in the MCF exhibit such a property: the amplitude distributions among cores within each ring adhere to a consistent pattern, aligning perfectly with the analytical predictions derived from Eq. (4.16). To bolster the validation of our theory, we specifically examine the amplitudes of two super-modes (modes 1 and 4) of a one-ring 3-core MCF shown in Fig. 4.3a and d, respectively. Since these modes are non-degenerate, their amplitudes should match exactly with those predicted by Eqs. (4.15–4.16). According to these equations, the expected amplitudes on the cores are (0.5456, 0.5456, 0.5456, 0.1887, 0.1887, 0.1887) and (−0.1887, −0.1887, −0.1887, 0.5456, 0.5456, 0.5456) for the two respective modes. The numerical outcomes presented in Fig. 4.3a and d indicate the core amplitudes as (1, 1, 1, 0.29, 0.29, 0.29) and (−0.35, −0.35, −0.35, 1, 1, 1) respectively. Upon normalizing these vectors, they transform to (0.5545, 0.5545, 0.5545, 0.1608, 0.1608, 0.1608) and (−0.1907, −0.1907, −0.1907, 0.5449, 0.5449, 0.5449), which are in close agreement with the analytical predictions, demonstrating the effectiveness of the proposed method.

Further calculations confirm that the results computed by Eqs. (4.15–4.16) also align with the results presented in Fig. 4.3b, c, e, f. It is noteworthy that since modes 2 and 3 (Fig. 4.3b, e) and modes 5 and 6 (Fig. 4.3c, f) are degenerated mode pairs, which possess non-unique Eigen vectors. Thus, they can be represented as the linear combination of the degenerated mode vectors computed by Eqs. (4.15–4.16).

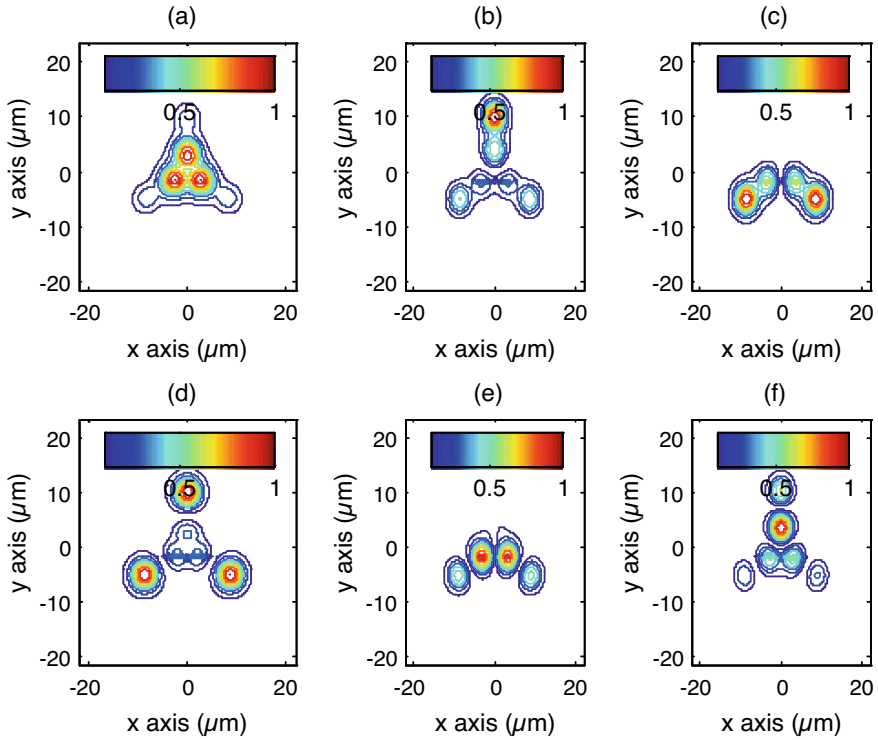


Fig. 4.3 The supermodes of the 6-core-MCF obtained by BPM (a–f) mode profiles of the supermodes 1–6. This figure is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

4.1.4 MCFs with Cores in Three Rings

Considering the MCFs with a three-ring core alignment, one may formulate the coupling matrix $\mathbf{M}$ as

$$\mathbf{M} = \begin{pmatrix} \kappa_1 & \mathbf{C}_{12} & 0 \\ \mathbf{C}_{21} & \kappa_2 & \mathbf{C}_{23} \\ 0 & \mathbf{C}_{32} & \kappa_3 \end{pmatrix} \quad (4.21)$$

where $\mathbf{C}_{ij}$ denotes the coupling between the i th ring and the j th ring and $\mathbf{C}_{ij} = \mathbf{C}_{ji}^H$. We may conduct the submatrix decomposition and obtain:

$$\mathbf{M} = \begin{pmatrix} \mathbf{Q} & & \\ & \mathbf{Q} & \\ & & \mathbf{Q} \end{pmatrix} \begin{pmatrix} \mathbf{D}_1 & \mathbf{C}_{12} & \mathbf{0} \\ \mathbf{C}_{21} & \mathbf{D}_2 & \mathbf{C}_{23} \\ \mathbf{0} & \mathbf{C}_{32} & \mathbf{D}_3 \end{pmatrix} \begin{pmatrix} \mathbf{Q}^H & & \\ & \mathbf{Q}^H & \\ & & \mathbf{Q}^H \end{pmatrix} \quad (4.22)$$

where $\mathbf{Q}$ is an orthogonal matrix and $\mathbf{D}_i$ a diagonal matrix. To derive the analytical expression for the Eigen values, we decompose the matrix in the middle as:

$$\begin{pmatrix} \mathbf{D}_1 & \mathbf{C}_{12} & \mathbf{0} \\ \mathbf{C}_{21} & \mathbf{D}_2 & \mathbf{C}_{23} \\ \mathbf{0} & \mathbf{C}_{32} & \mathbf{D}_3 \end{pmatrix} = \mathbf{U}\mathbf{N}\mathbf{U}^H \quad (4.23)$$

where $\mathbf{N}$ is a block matrix

$$\mathbf{N} = \begin{pmatrix} d_1^1 & c_{12} & & & \\ c_{21} & d_2^1 & c_{23} & & \\ & c_{32} & d_3^1 & & \\ & & & \ddots & \\ & & & & d_1^N & c_{12} \\ & & & & c_{21} & d_2^N & c_{23} \\ & & & & & c_{32} & d_3^N \end{pmatrix} \quad (4.24)$$

and the matrix $\mathbf{U}$ is a matrix that arises from the permutation of rows within the identity matrix. The k th row of the identity matrix is exchanged with the $(3k-2)$ th row when k is between $[1, N]$, with the $(3(k-N)-1)$ th row when k is between $[N+1, 2N]$ or with the $3(k-2N)$ th row when k is between $[2N+1, 3N]$.

Analogous to the two-ring scenario, the eigenvalues and eigenvectors of the block matrix $\mathbf{N}$ are interconnected with the sub-matrices:

$$\begin{pmatrix} d_1^n & c_{12} \\ c_{21} & d_2^n & c_{23} \\ & c_{32} & d_3^n \end{pmatrix} \quad (4.25)$$

Since $c_{ij} = c_{ji}$, the equation to compute the eigen value is as follows:

$$\begin{aligned} \lambda^3 + A\lambda^2 + B\lambda + C &= 0 \\ A &= -(d_1^n + d_2^n + d_3^n) \\ B &= (d_1^n d_2^n + d_1^n d_3^n + d_2^n d_3^n - c_{12}^2 - c_{23}^2) \\ C &= -d_1^n d_2^n d_3^n + c_{23}^2 (d_1^n + d_3^n) \end{aligned} \quad (4.26)$$

There are three roots for the above Eq. (4.26), and they can be calculated as [27]

$$\lambda_k = -\frac{1}{3} \left(b + u_k D + \frac{\Delta_0}{u_k D} \right) \quad (4.27)$$

where [27]

$$\begin{aligned}
u_1 &= 1, u_2 = \frac{-1 + j\sqrt{3}}{2}, u_3 = \frac{-1 - j\sqrt{3}}{2} \\
D &= \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}} \\
\Delta_0 &= B^2 - 3c \\
\Delta_1 &= 2B^3 - 9BC + 27d
\end{aligned} \tag{4.28}$$

Having determined the eigenvalues, the eigenvectors of the sub-matrices can also be derived analytically. Consequently, the propagation constants and the Eigen mode vectors can be acquired in a manner like the two-ring case. These new eigenvectors remain as superpositions of the eigenvectors of the coupling matrices of the individual rings.

4.1.5 MCFs with Cores in L Rings

4.1.5.1 Mathematical Formulation

We may extend the analysis to the MCFs with cores distributed in L rings, and obtain the coupling matrix as

$$\mathbf{M} = \begin{pmatrix} \kappa_1 & \mathbf{C}_{12} & & \\ \mathbf{C}_{21} & \kappa_2 & \mathbf{C}_{23} & \\ & & \ddots & \\ & & & \mathbf{C}_{L,L-1} & \kappa_L \end{pmatrix} \tag{4.29}$$

where $\mathbf{C}_{ij}$ denotes the coupling matrix between the i th ring and the j th ring and $\mathbf{C}_{ij} = \mathbf{C}_{ji}^H$. Matrix decomposition can be applied to the matrix $\mathbf{M}$ as

$$\mathbf{M} = \mathbf{U}\mathbf{N}\mathbf{U}^H \tag{4.30}$$

where $\mathbf{N}$ is a block matrix constructed by sub-matrices of the form

$$\begin{pmatrix} d_1^n & c_{12} & & \\ c_{21} & d_2^n & c_{23} & \\ & & \ddots & \\ & & & c_{L,L-1} & d_L^n \end{pmatrix} \tag{4.31}$$

where $c_{ij} = c_{ji}$. The aforementioned matrix is Hermitian or symmetric, which possesses real eigenvalues. Additionally, it is a band matrix characterized by a

polynomial of degree L . Efficient numerical techniques can be employed to find its eigenvalues and eigenvectors, which is a topic that has gained significant attention in numerous publications [28].

The final propagation constants can be computed by adding β to the eigenvalues obtained from Eq. (4.31). Furthermore, the amplitude distribution vectors of the supermodes are formulated as the linear combination of the eigenvectors calculated from the coupling matrices of the individual rings. The coefficients for this combination are evaluated by the eigenvectors of the sub-matrix defined in Eq. (4.31).

4.1.5.2 A Special Case When the Radiuses of the Rings Are Very Large

With the large radii rings, the coupling strength can be regarded as uniform. In this case, the coupling matrix in Eq. (4.31) should take the form:

$$\begin{pmatrix} 2\kappa \cos\left(\frac{2\pi(n-1)}{N}\right) & \kappa_a & & \\ \kappa_a & 2\kappa \cos\left(\frac{2\pi(n-1)}{N}\right) & \kappa_a & \\ & & \ddots & \\ & & & \kappa_a & 2\kappa \cos\left(\frac{2\pi(n-1)}{N}\right) \end{pmatrix} \quad (4.32)$$

where κ denotes the coupling coefficient inside individual rings and κ_a signifies the coupling coefficient between neighboring rings. It is quite straight forward to compute the m th Eigen value of the matrix as $2\kappa \cos\left(\frac{2\pi(n-1)}{N}\right) + 2\kappa_a \cos\left(\frac{\pi m}{L+1}\right)$ and the respective Eigen vectors as

$$\sqrt{\frac{2}{L+1}} \left(\sin\left(\frac{\pi m}{L+1}\right), \sin\left(\frac{\pi 2m}{L+1}\right), \dots, \sin\left(\frac{\pi lm}{L+1}\right) \right)^T \quad (4.33)$$

This indicates that the modal amplitude distribution can be separated into two distributions in the radial and the angular directions. It adheres to a sinusoidal function in the radial direction, whereas the angular direction amplitude distribution follows the pattern described in Eqs. (4.4–4.5).

4.1.6 MCFs with a Core in the Middle of the Rings

4.1.6.1 Mathematical Formulation

As is mentioned in the introduction, the MCF can be applied with a center core to increase the transmission efficiency. In this case, the coupling matrix should be modified as:

$$\mathbf{M} = \begin{pmatrix} 0 & \mathbf{b}^T & & & \\ \mathbf{b} & \kappa_1 & \mathbf{C}_{12} & & \\ & \mathbf{C}_{21} & \kappa_2 & \mathbf{C}_{23} & \\ & & & \ddots & \\ & & & & \mathbf{C}_{L-1,L} & \kappa_L \end{pmatrix} \quad (4.34)$$

where $\mathbf{b} = b_0(1,1,\dots,1)^T$ describes the coupling between the center core and the cores within the first ring.

As is shown in [24], the following matrix

$$\mathbf{M}_1 = \begin{pmatrix} 0 & \mathbf{b}^T \\ \mathbf{b} & \mathbf{Q}\mathbf{D}_1\mathbf{Q}^H \end{pmatrix} \quad (4.35)$$

has $N + 1$ Eigen values, which can be calculated directly. Among them, $N-1$ Eigen values are the same as matrix κ_I , while the two more Eigen values can be computed analytically [24]

$$\kappa_1 \pm \sqrt{\kappa_1^2 + Nb_0^2} \quad (4.36)$$

The Eigen vectors can be evaluated analytically as well. $N-1$ Eigen vectors are the same as those of matrix κ_I except an additional 0 element, because the size of the vector is $N + 1$. There are two more Eigen vectors, which can be calculated as [24]

$$\left(\frac{-\kappa_1 \pm \sqrt{\kappa_1^2 + Nb_0^2}}{b_0}, 1, 1, \dots, 1 \right)^T \quad (4.37)$$

These Eigen vectors will form the two columns of the orthogonal matrix $\mathbf{Q}'$ and the matrix decomposition $\mathbf{M}_1 = \mathbf{Q}'\mathbf{D}_1\mathbf{Q}'^H$ can be applied.

Thereafter, we may apply the matrix decomposition for the matrix $\mathbf{M}$ in Eq. (4.34) as

$$\begin{aligned}
\mathbf{M} &= \begin{pmatrix} \mathbf{Q}'\mathbf{D}_1\mathbf{Q}^H & \mathbf{Q}'\mathbf{C}'^H\mathbf{Q}^H & & \\ \mathbf{Q}\mathbf{C}'\mathbf{Q}^H & \mathbf{Q}\mathbf{D}_2\mathbf{Q}^H & \mathbf{Q}\mathbf{C}_{23}\mathbf{Q}^H & \\ & & \ddots & \\ & \mathbf{Q}\mathbf{C}_{L-1,L-2}\mathbf{Q}^H & \mathbf{Q}\mathbf{D}_{L-1}\mathbf{Q}^H & \mathbf{Q}\mathbf{C}_{L-1,L}\mathbf{Q}^H \\ & & \mathbf{Q}\mathbf{C}_{L,L-1}\mathbf{Q}^H & \mathbf{Q}\mathbf{D}_L\mathbf{Q}^H \end{pmatrix} \\
&= \begin{pmatrix} \mathbf{Q}' & & & \\ & \mathbf{Q} & & \\ & & \ddots & \\ & & & \mathbf{Q} \end{pmatrix} \begin{pmatrix} \mathbf{D}_1 & \mathbf{C}'^H & & \\ \mathbf{C}' & \mathbf{D}_2 & \mathbf{C}_{23} & \\ & & \ddots & \\ & \mathbf{C}_{L-1,L-2} & \mathbf{D}_{L-1} & \mathbf{C}_{L-1,L} \\ & & \mathbf{C}_{L,L-1} & \mathbf{D}_L \end{pmatrix} \begin{pmatrix} \mathbf{Q}^H & & & \\ & \mathbf{Q}^H & & \\ & & \ddots & \\ & & & \mathbf{Q}^H \end{pmatrix} \quad (4.38)
\end{aligned}$$

where $\mathbf{Q}$ represents the matrix whose rows consist of the eigenvectors of the individual rings' coupling matrix, and $\mathbf{Q}'$ denotes the matrix where its rows are formed by the eigenvectors of the coupling matrix between the center core and the cores located within the first ring, specifically referring to matrix $\mathbf{M}_1$. Matrix $\mathbf{C}'$ is

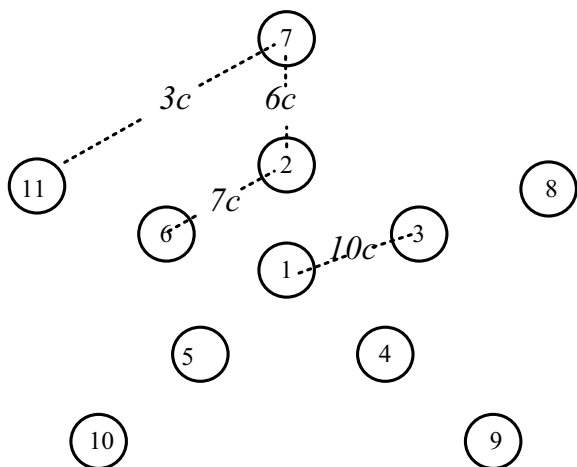
$$\begin{aligned}
\mathbf{C}' &= \mathbf{Q}^H \begin{pmatrix} 0 & c_{1,2} & & \\ \vdots & & \ddots & \\ 0 & & & c_{1,2} \end{pmatrix} \mathbf{Q}' = \begin{pmatrix} c'_+ & c'_- & & \\ \vdots & & \ddots & \\ 0 & & & c_{1,2} \end{pmatrix} \quad (4.39) \\
c'_{\pm} &= c_{1,2} \sqrt{\frac{N}{N + \left(\frac{-\kappa_1 \pm \sqrt{\kappa_1^2 + Nb_0^2}}{b_0} \right)^2}}
\end{aligned}$$

Further simplification of the matrix can be done by decomposing the matrix in the center of Eq. (4.38) into $\mathbf{U}\mathbf{N}\mathbf{U}^H$, where the matrix $\mathbf{N}$ is made of block matrices, which is formulated by Eq. (4.31) for $n = 2 \dots N$ with the first block matrix as.

$$\begin{pmatrix} d_{1-}^1 & c'_- & & \\ & d_{1+}^1 & c'_+ & \\ c'_- & c'_+ & d_2^1 & \\ & & & \ddots \\ & c_{L-1,L-2} & d_{L-1}^1 & c_{L-1,L}^1 \\ & & c_{L,L-1} & d_L^1 \end{pmatrix} \quad (4.40)$$

where the first two diagonal elements in Eq. (4.40) can be formulated by Eq. (4.36). It can be inferred from Eq. (4.40) that the last $(N-1)L$ Eigen values of the matrix $\mathbf{N}$ is consistent with the those in Eq. (4.31) and their corresponding eigenvectors also remain the same, except for the inclusion of an additional zero component. The rest $L + 1$ Eigen values should be calculated by evaluating the Eigen values of the matrix

Fig. 4.4 The 11-core-MCF core array arrangement and the assumed coupling coefficients between the cores. This figure is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]



in Eq. (4.40). Their corresponding eigenvectors also remain the same, except for the inclusion of an additional zero component.

4.1.6.2 An Example

The proposed analytical formulas are verified by numerical simulations with a MCF of two rings and a center core. The MCF has eleven cores, with ten of them in the two rings and one of them in the center. The first ring cores are assumed with the coupling coefficient of $7c$ and the second ring cores are assumed with the coupling coefficient of $3c$. The adjacent cores of the different rings are assumed with the coupling coefficient of $6c$. The center core and the cores of the first ring is assumed with the coupling coefficient of $10c$. The optical schematic is shown in Fig. 4.4.

To substantiate the claim that a multi-core fiber (MCF) featuring a central core within its ring structure and an MCF without such a central core share $(N-1)L$ same eigenvalues, we further analyze the MCF resembling the one depicted in Fig. 4.4, except that it lacks a central core within its configuration.

The eigenvalues are straightforwardly calculated from the expressions of the two coupling matrices associated with the MCFs, utilizing a numerical approach, and the results are shown in Table 4.2. It is evident from the table that the two matrices share numerous identical eigenvalues, with the exception of three unique eigenvalues belonging solely to the first matrix. This observation aligns perfectly with the aforementioned conclusion. Additionally, the eigenvalues can be derived using Eq. (4.14), while the three distinctive eigenvalues can be precisely calculated through Eq. (4.40), yielding results that match those presented in Table 4.2.

Table 4.2 Eigen values of the two coupling matrices of the MCFs with and without a center core

Eigen values of the coupling matrix with a center core	Eigen values of the coupling matrix of the two ring MCF
−16.9918c	
−14.9072c	−14.9072c
−14.9072c	−14.9072c
−3.0358c	−3.0358c
−3.0358c	−3.0358c
−1.2731c	−1.2731c
−1.2731c	−1.2731c
5.6296c	2.7889c
9.2162c	9.2162c
9.2162c	9.2162c
31.3622c	17.2111c

This table is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

4.1.7 Case of Strong Coupling Among the Cores

The analytical framework employed in this study presupposes that coupling solely occurs between immediately adjacent cores, a widely adopted approach in analyzing the super-modes of MCFs [15–17, 19, 22–24]. However, in scenarios where cores are arranged in extremely close range, coupling may extend beyond immediate neighbors. Additionally, the coupled mode theory maintains precision in the scenarios of weak to moderate coupling, and it will experience diminished accuracy under conditions of strong coupling [26]. To study the deviations originated from these considerations, a comparative analysis can be conducted on the outcomes derived from analytical formulations and those from numerical simulations that comprehensively account for all coupling effects [29].

To study the influence of non-neighboring core coupling and strong coupling, a comprehensive analysis of a six-core MCF with close core arrangement is undertaken. This fiber maintains the same cladding and core refractive indices as those discussed in Sect. 4.1.3, with core diameters of 5 μm and inter-core distances of 5.5 μm , resulting in short 0.5 μm gaps between neighboring cores. A diamond lattice configuration is assumed for the MCF. Under the assumption of neighboring core coupling, the effective indices of the super-modes can be approximately estimated using Eq. (4.3). A comparison is made between numerical simulations conducted using the BPM and the theoretical estimates, as depicted in Fig. 4.5. This figure plots the effective indices of both the 1st and 2nd order modes as functions of wavelength, clearly showing the theory's accuracy relative to the numerical results. However, as the gap size diminishes further, for instance, reaching zero gap size, the inaccuracy of the theory becomes notably pronounced, as evidenced in Fig. 4.6. To precisely predict

the effective indices in such tightly packed configurations, a more refined analytical theory is imperative, which will be presented in the latter section.

To gain a deeper understanding of the effects of coupling between non-adjacent cores, a four-core MCF with a square lattice arrangement is examined. This configuration, featuring four cores arranged in a ring, inherently exhibits the strongest non-adjacent core coupling compared to alternative structures. The four-core MCF employed in the simulations shares the same core/cladding refractive indices and core diameters as the six-core MCF discussed earlier. Initially, a gap size of $0.5\ \mu\text{m}$ is assumed between the cores.

Table 4.3 presents a comparison of the effective indices obtained using both analytical and numerical methods, specifically at the signal wavelength of $1550\ \text{nm}$. It becomes evident that the discrepancy between the two methods increases when

Fig. 4.5 The effective indexes of the super-modes of the six-core MCF. (gap size = $0.5\ \mu\text{m}$). This figure is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

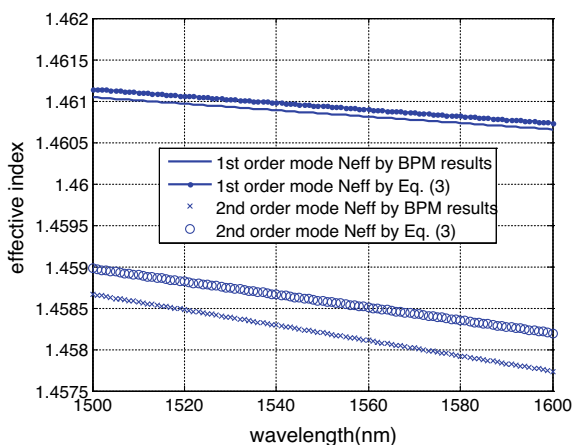


Fig. 4.6 The effective indexes of the super-modes of the six-core MCF. (gap size = $0\ \mu\text{m}$). This figure is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

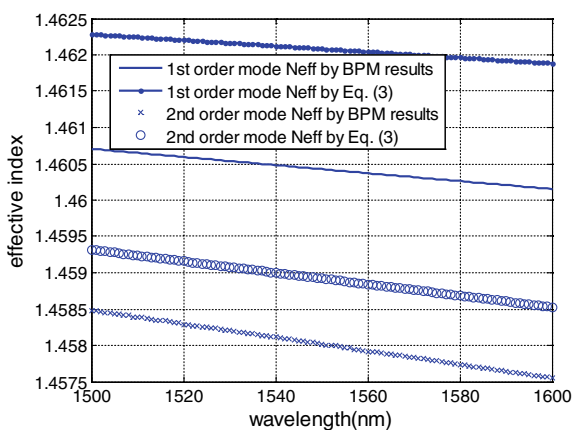


Table 4.3 The effective indexes of the super-modes of the four-core MCF (gap size = 0.5 μm)

Mode number	BPM results	Results from Eq. (4.3)
1	1.4610	1.4613
2	1.4560	1.4566
3	1.4560	1.4566
4	1.4515	1.4519

This table is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

Table 4.4 The effective indexes of the super-modes of the four-core MCF (gap size = 0 μm)

Mode number	BPM results	Results from Eq. (4.3)
1	1.4617	1.4608
2	1.4556	1.4566
3	1.4556	1.4566
4	1.4524	1.4524

This table is reprinted/adapted with permission from [Opt. Express 22, 673–688 (2014)] © The Optical Society of America [21]

compared to the six-core case, highlighting the enhanced influence of non-adjacent coupling in the square lattice.

Furthermore, as the gap size is reduced to zero, as is shown in Table 4.4, the inaccuracy of the analytical approach escalates significantly. To maintain an effective index error margin below 0.3%, it is crucial to ensure that the gap between the cores remains no less than 0.7 μm . This threshold underscores the significance of core spacing in the accuracy of the theory under the strong and non-adjacent core coupling.

4.1.8 Summary

We have conducted a comprehensive analysis of multi-core fibers (MCFs) featuring circularly arranged cores. For these configurations, we have derived analytical expressions encompassing two-ring, three-ring, and even multi-ring scenarios. Additionally, we have studied the specific case where an extra core is positioned at the center of the circular fiber array. The expressions derived in this section will undoubtedly benefit those engaged in the design and analysis of MCFs, providing valuable insights and facilitating their work. The theory will be extended to the strong coupling case in the next section.

4.2 A Non-orthogonal Coupled Mode Theory for Super-Modes Inside Multi-core Fibers

Part of the section reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30].

In the previous section, the coupled mode theory assumes that the modes of the individual cores have little overlap and therefore they are orthogonal to each other. Such assumption is not valid when the cores are close to each other and the coupling becomes strong. In this section, a non-orthogonal coupled mode theory is proposed to analyze the super-modes of multi-core fibers (MCFs), which is valid in the strong coupling regime. MCFs with circularly distributed cores are analyzed as an example. Analytical formulas are derived both for the refractive indexes and the Eigen vectors of the super-modes. It is rigorously revealed that the Eigen vectors for the super-modes of such MCFs are the row vectors of the inverse discrete Fourier transform (IDFT) matrix. Therefore, by pre-coding the signal channels via IDFT, one is able to generate the super-modes for the MCFs with circularly distributed cores.

4.2.1 Introduction

As is shown in the previous section, MCFs are of great importance in signal transmission, lasing and other applications. Therefore, quite a few researches have been devoted to the study of MCFs, particularly on the property of the super-modes. Since these progresses have been discussed thoroughly in the previous section, it will not be reiterated here. As is demonstrated, most of the studies only tackle the MCFs with weak coupling.

In conjunction with the study of super-modes residing within MCFs, methodologies for the multiplexing of these super-modes have emerged. The mode index matching method by Saitoh et al. [10] enabled mode conversion from a single mode core to the super-mode with similar refractive index. Arrayed waveguide gratings (AWGs) were proposed by Kokubun et al [11] to function as mode multiplexers. Zhou et al. [31] proposed that the MCF mode multiplexers can be achieved by proper design of multimode interference (MMI) couplers. While the pioneering works are abundant, they are mostly suitable for the MCFs with linearly arranged cores [10, 11, 31] and there have few reports on the mode multiplexers designed for MCFs with circularly arranged cores.

In this section, a rigorous non-orthogonal CMT [32, 33] is proposed to analyze the MCFs in the strong coupling regime. As an illustrative case, we conduct the study on the MCFs featuring circularly arranged cores. We derive the analytical expressions for the refractive index and eigenmodes of the super-modes in the circularly aligned MCFs, ensuring their consistency with established theories. Our findings reveal that the super-mode vectors strictly conform to the row vectors of the IDFT matrix, even under the conditions of strong coupling between the cores. To substantiate these

analytical formulations, we conduct numerical simulations, which exhibit remarkable accordance with our theoretical predictions.

Apart from the super-mode analysis, super-mode multiplexing technique based on the IDFT algorithm is proposed. The signal channel pre-coding technique require no MCF super-modes MUXs, which can greatly simplify the system architecture.

4.2.2 A Rigorous Non-Orthogonal Coupled Mode Theory

The most important problem for the conventional CMT [16, 19, 22, 24] is that it assumes that the cross-power term and the coupling between non-neighboring cores are zero. Such simplification induces the inaccuracy while predicting the super-modes [21, 26, 32, 33] in the MCF, particularly in the strongly coupled regime [21, 32, 33]. In this context, we aim to derive a theory based on the non-orthogonal CMT [32], which is given satisfying predictions in the strong coupling regime.

To study the super-modes of the MCF, we start with the scalar Holmholtz equation describing the field evolution inside the MCFs [29]

$$\nabla^2 \varphi + k^2 \varphi = 0 \quad (4.41)$$

The field inside the waveguides can be expressed by combination of the modes of the MCF cores [29]

$$\varphi = \sum_{n=1}^N c_n \varphi_n e^{j\beta z} \quad (4.42)$$

where φ_n stands for the fundamental mode of the n th core, c_n represents the coefficient of the corresponding mode, β denotes the propagation constant of the respective mode. Each fundamental mode should fulfill the mode equation in case other cores are not present [29]:

$$\nabla_{\perp}^2 \varphi_n + (k_n^2 - \beta^2) \varphi_n = 0 \quad (4.43)$$

where k_n denotes the free space wavelength number timed by the refractive indices of each core and the related cladding, the symbol $\nabla_{\perp}^2$ represents the two-dimensional Laplace operator on the plane of the MCF cross section.

Combining Eqs. (4.41–4.43) and ignoring the second order derivative with respect to z , one may derive the following equation:

$$\sum_{n=1}^N \frac{\partial c_n}{\partial z} \varphi_n = \sum_{n=1}^N c_n \frac{j(k^2 - k_n^2)}{2\beta} \varphi_n \quad (4.44)$$

Without loss of generality, φ_n may be assumed to be real [18]. Equation (4.44) can be converted into the coupled mode equation by the integration over the cross section of the plane [32]

$$\mathbf{S} \frac{\partial \mathbf{c}}{\partial z} = j \mathbf{E} \mathbf{c} \quad (4.45)$$

where vector $\mathbf{c}$ is composed of the core mode amplitudes and the two matrixes $\mathbf{E}$ and $\mathbf{S}$ represent the interaction between the cores:

$$\begin{aligned} E_{mn} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{(k^2 - k_n^2)}{2\beta} \varphi_m \varphi_n dx dy \\ S_{mn} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi_m \varphi_n dx dy \end{aligned} \quad (4.46)$$

The derived non-orthogonal coupled mode theory can be implemented on the MCFs with strong coupling and the analytical expressions can be derived based on it.

4.2.3 Analytical Formulation of Super-Modes Inside Circularly Aligned MCFs

As is mentioned earlier, we consider the MCFs with N cores with equal distance and is arranged on the circular ring. In this case, the elements of the matrices should have the following property:

$$\begin{aligned} E_{mn} &= E_{m'n'} \\ S_{mn} &= S_{m'n'} \\ |m - n| &= |m' - n'| \end{aligned} \quad (4.47)$$

Henceforth, we have only N independent variables within the matrix and we may denote them as $E_1 \dots E_N$, and $S_1 \dots S_N$, which are:

$$\begin{aligned} E_{mn} &= E_{|m-n|+1} \\ S_{mn} &= S_{|m-n|+1} \end{aligned} \quad (4.48)$$

Therefore, the matrices $\mathbf{E}$ and $\mathbf{S}$ are circulant matrices. According to their mathematical property, the matrices can be formulated as [21, 34]

$$\begin{aligned}\mathbf{E} &= \mathbf{Q}\mathbf{G}\mathbf{Q}^H \\ \mathbf{S} &= \mathbf{Q}\mathbf{D}\mathbf{Q}^H\end{aligned}\tag{4.49}$$

where $\mathbf{Q}$ is the IDFT matrix:

$$Q_{mn} = \frac{1}{\sqrt{N}} \exp\left(j \frac{2\pi(m-1)(n-1)}{N}\right)\tag{4.50}$$

and $\mathbf{Q}^H$ is the discrete Fourier transform (DFT) matrix. The analytical expression for the diagonal elements of the two diagonal matrixes $\mathbf{G}$ and $\mathbf{D}$ should be [34]

$$\begin{aligned}G_{nn} &= \sum_{m=1}^N E_m \exp\left(-j \frac{2\pi(n-1)(m-1)}{N}\right) \\ D_{nn} &= \sum_{m=1}^N S_m \exp\left(-j \frac{2\pi(n-1)(m-1)}{N}\right)\end{aligned}\tag{4.51}$$

Therefore, one may recast Eq. (4.45) into:

$$\frac{\partial \mathbf{c}}{\partial z} = j\mathbf{M}\mathbf{c}\tag{4.52}$$

where

$$\mathbf{M} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{G}\mathbf{Q}^H\tag{4.53}$$

Hence, the row vectors of $\mathbf{Q}$ actually form super-mode vectors and the corresponding propagation constants of these super-modes can be obtained analytically as

$$\beta_n = \beta + \frac{G_{nn}}{D_{nn}}\tag{4.54}$$

Since circular ring is symmetric, we have.

$$\begin{aligned}E_m &= E_{N-m} \\ S_m &= S_{N-m}\end{aligned}\tag{4.55}$$

Therefore, one has.

$$\beta_n = \beta_{N-n+2}\tag{4.56}$$

This implies that mode n and mode $N-n+2$ are degenerate modes, resulting in non-uniqueness of the matrix $\mathbf{Q}$. To convert the complex amplitudes of the super-modes into real values, a particular form of $\mathbf{Q}$ has been suggested in [21], as is shown in the following expression:

If N is even

$$Q_{mn} = \begin{cases} \sqrt{\frac{1}{N}} & n = 1 \\ \sqrt{\frac{2}{N}} \cos((m-1)(n-1)\frac{2\pi}{N}) & 1 < n \leq \frac{N}{2} \\ \sqrt{\frac{1}{N}} (-1)^{m-1} & n = 1 + \frac{N}{2} \\ \sqrt{\frac{2}{N}} \sin((m-1)(N-n+1)\frac{2\pi}{N}) & \frac{N}{2} + 1 < n \leq N \end{cases} \quad (4.57a)$$

If N is odd

$$Q_{mn} = \begin{cases} \sqrt{\frac{1}{N}} & n = 1 \\ \sqrt{\frac{2}{N}} \cos((m-1)(n-1)\frac{2\pi}{N}) & 1 < n \leq \frac{N+1}{2} \\ \sqrt{\frac{2}{N}} \sin((m-1)(N-n+1)\frac{2\pi}{N}) & \frac{N+1}{2} < n \leq N \end{cases} \quad (4.57b)$$

4.2.4 Comparison with the Published Results

The correctness of the proposed theory is tested by comparing with the published results. For a two-core MCF, the two super-mode modes should have their effective indices according to the theory:

$$\begin{aligned} n_{eff1} &= \frac{\beta_1}{k_0} = n_{eff0} + \frac{k_{11} + k_{12} - X(k_{11} + k_{12})}{1 - X^2} \\ n_{eff2} &= \frac{\beta_2}{k_0} = n_{eff0} + \frac{k_{11} - k_{12} + X(k_{11} - k_{12})}{1 - X^2} \end{aligned} \quad (4.58)$$

where

$$\begin{aligned} k_{11} &= \frac{E_1}{k_0 S_1} \\ k_{12} &= \frac{E_2}{k_0 S_1} \\ X &= \frac{S_2}{S_1} \\ n_{eff0} &= \frac{\beta}{k_0} \end{aligned}$$

The same expression as Eq. (4.58) has been derived in [33] (Eq. (2.1a–2.1b)), clearly demonstrating the correctness of the propose theory.

4.2.5 Numerical Simulations and Discussions

In addition to the comparison with the published results, the formulas are verified through numerical simulations based on the three-dimensional beam propagation method (3DBPM). A closely aligned six core MCF is used as an example, with its cross section shown in the following Fig. 4.7.

The cladding index is specified as 1.45, while the core exhibits an index difference of 1.2% relative to the cladding, with the radii of $2.5\ \mu\text{m}$. These index values and radii are the same as those in [10] to ensure the example's validity and realism. The cores are positioned with a center-to-center distance of $5\ \mu\text{m}$, resulting in a close alignment without any gaps between them. As is demonstrated in [21], this particular core arrangement leads to a pronounced coupling effect among the cores, which introduces remarkable inaccuracies in the evaluation of the super-modes' effective indices using conventional coupled-mode theory (CMT).

Table 4.5 presents a comparative analysis of the effective refractive indices of super-modes, computed using three distinct methods: the 3D Beam Propagation Method (3DBPM), Eq. (4.54), and the simplified conventional Coupled Mode Theory (CMT). The comparison underscores the validity of the theoretical framework and highlights the substantial enhancement in accuracy achieved by the non-orthogonal CMT over the simplified CMT, which solely considers the nearest-neighbor coupling effect.

Fig. 4.7 Cross section of the six core fiber. This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

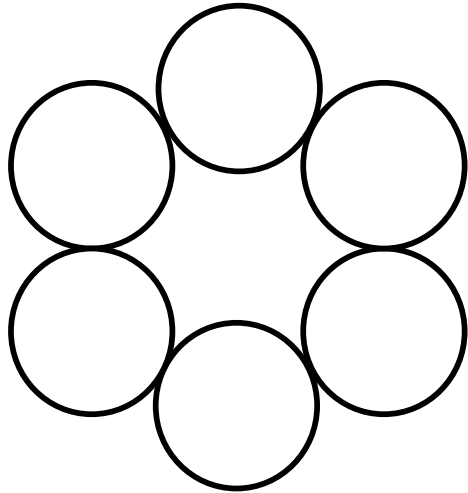


Table 4.5 The calculated refractive indexes of the super-modes by the BPM, Eq. (4.44) and the conventional CMT

Super-mode Number	BPM results	Equation (4.14)	Conventional CMT
1	1.4619	1.4619	1.4620
2	1.4607	1.4607	1.4604
3	1.4579	1.4575	1.4571
4	1.4554	1.4552	1.4554
5	1.4579	1.4575	1.4571
6	1.4607	1.4607	1.4604

This table is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

Subsequently, the super-mode fields derived from the Beam Propagation Method (BPM) simulations and analytical expressions are shown in Figs. 4.8 and 4.9. Given that BPM simulations implement super-mode fields with uniform phases across the x–y plane, Eq. (4.57) is employed instead of Eq. (4.50) for comparison. The correlation coefficients between the mode pairs depicted in Figs. 4.8 and 4.9, defined as normalized overlap integrals, are calculated to be 1, 0.99, 0.98, 0.99, 0.93, and 0.92, respectively. This signifies an excellent agreement between the results presented in Figs. 4.8 and 4.9.

In addition to the six-core MCF, one more example of a four-core MCF is tested with the theory. The radii of the cores remain the same as 2.5 μm , and the cladding index maintains 1.45. But the index difference is reduced to 0.6%, which might result in more significant coupling effect. The effective refractive indices of the three super-modes were computed using both BPM simulations and Eq. (4.54), with the results shown in Table 4.6. It is evident that the values obtained from Eq. (4.54) align well with the BPM outcomes. The reason behind the four-core MCF exhibiting only three super-modes lies in the significant modal field interaction that occurs when the phase difference between neighboring cores reaches π , which might diminish the modal field. Specifically, a mode with such a phase difference cannot exist within the four-core MCF, as confirmed by finite element method (FEM) simulations. Notably, even in this extreme scenario, the formulas provide an accurate estimation of the effective refractive indices for the existing super-modes.

4.2.6 Signal Pre-Coding Based on the Concept

It is interesting to note that the proposed theory can not only provide a theoretical prediction for the effective refractive and mode distribution for the super-modes, it can also be used for super-mode generation and multiplexing. One may note that the eigen vectors of the super-modes are actually the row vectors of the IDFT matrix. In this way, if the transmitted data can be processed by the IDFT, the signal vector $\mathbf{g}$

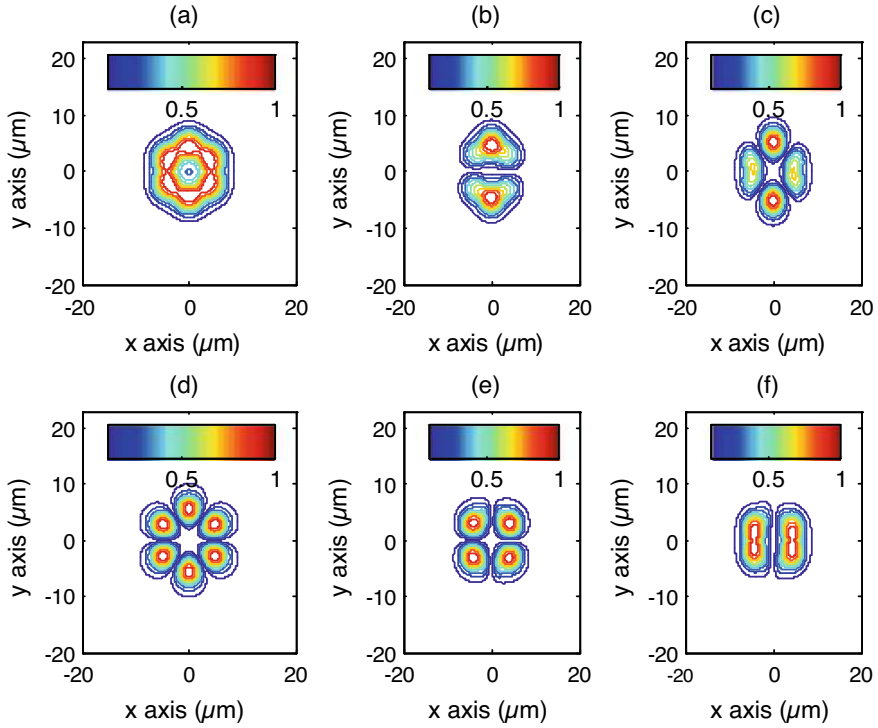


Fig. 4.8 Mode fields calculated by BPM simulations. This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

can be converted to:

$$\mathbf{Qg} \quad (4.59)$$

The data transmits inside the MCF, and evolves into the following at the propagation distance of L :

$$\mathbf{QPQ}^H \mathbf{Qg} = \mathbf{QPg} \quad (4.60)$$

where $\mathbf{P}$ is a diagonal matrix and can be formulated as:

$$\mathbf{P} = \text{diag}(\exp(j\beta_1 L), \exp(j\beta_2 L), \dots, \exp(j\beta_N L)) \quad (4.61)$$

where β_n is the propagation constant of each super-mode. In this way, IDFT has multiplexed the optical signal on the super-modes of the MCF without any actual optical devices as is shown in Fig. 4.10.

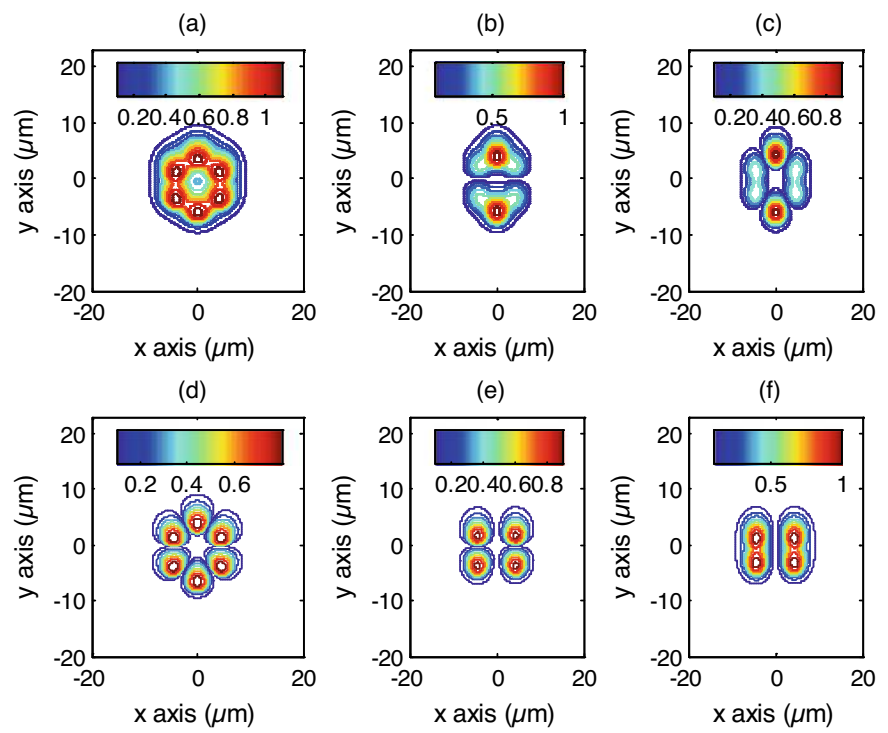


Fig. 4.9 Mode fields calculated by Eq. (4.17). This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

Table 4.6 The calculated refractive indexes of the super-modes of a four core MCF by the BPM, Eq. (4.44) and the conventional CMT

Mode number	BPM results	Equation (4.14)	Conventional CMT
1	1.4550	1.4548	1.4558
2	1.4522	1.4520	1.4524
3	1.4522	1.4520	1.4524

This table is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

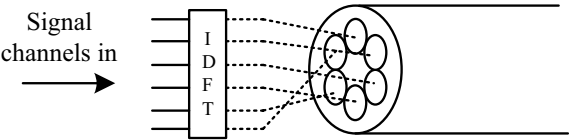


Fig. 4.10 IDFT based super-mode multiplexing technique. This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

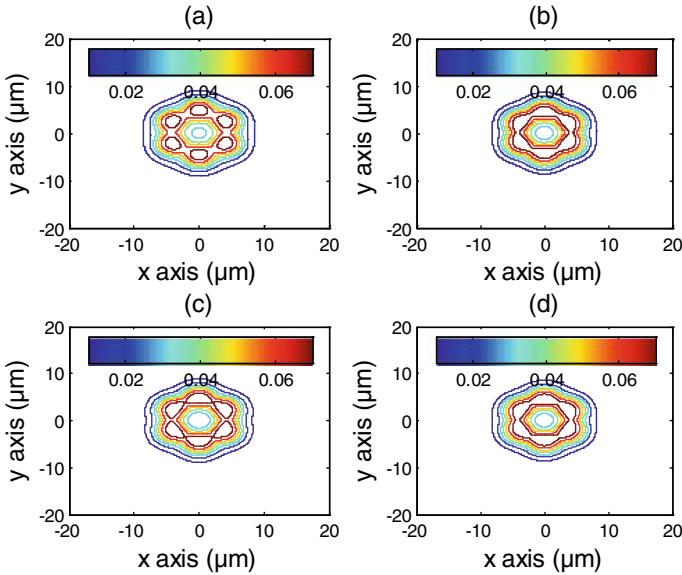


Fig. 4.11 The fields of the MCF with the input vector as $[1 \ 0 \ 0 \ 0 \ 0 \ 0]$ before pre-coding **a** at the input of the MCF, **b** after propagating inside the MCF for 1 m, **c** after propagating inside the MCF for 5 m, **d** after propagating inside the MCF for 10 m. This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

The multiplexing techniques are exemplified as follows. We consider the six-core MCF previously discussed for simulation purposes, operating at the signal wavelength of 1550 nm. The input vector is presumed to be $[1 \ 0 \ 0 \ 0 \ 0 \ 0]$, representing activation of only the first core. Upon application of the IDFT, this vector transforms into the coefficients of the fundamental modes within each core.

The initial input signal field configuration is depicted in Fig. 4.11a, while the evolved output signals after propagating through the fiber for 1 m, 5 m, and 10 m distances are presented in Fig. 4.11b–d, respectively. From Fig. 4.11, it becomes evident that both the input and the output signals effectively function as the fundamental supermode of the MCF, retaining their essential characteristics even after propagation. This observation underscores the validity and efficacy of the proposed pre-coding concept.

4.2.7 De-Multiplexing the Signals Based on IDFT

For signal de-multiplexing, a tapered fiber design can be employed, which gradually disentangles the embedded cores within the MCF. Assuming the core-to-core distances remain uniform, as discussed previously, the supermode vectors will persist

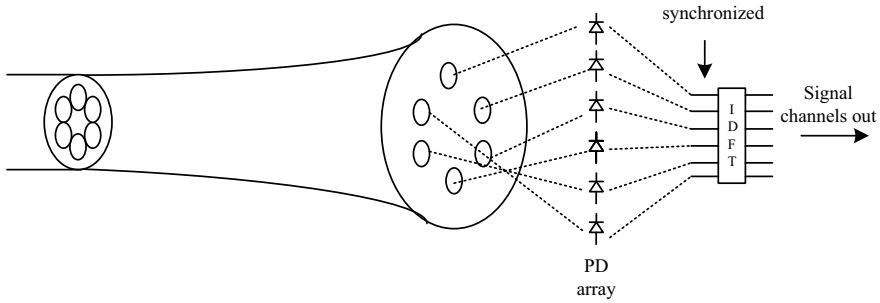


Fig. 4.12 IDFT based super-mode de-multiplexing technique. This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

as the row vectors of the IDFT matrix. Consequently, the supermode information is preserved during the core separation process.

Once the cores are sufficiently spaced apart, the information carried by each core can be individually captured by a photodiode (PD) array. In cases where coherent detection is utilized, not only the amplitudes but also the phases of the modes on each core can be accurately detected. Subsequently, by applying the IDFT to the detected signals, the original input signal can be faithfully reconstructed. The schematic outlining this de-multiplexing approach is depicted in Fig. 4.12.

To validate the proposed concept, a simulation was conducted for a tapered fiber with a tapering length of 10000 μm . The taper design maintains consistent core and cladding refractive indices throughout. At the input end, the cores are closely packed, whereas at the output end, adjacent cores are separated by a distance of 20 μm . Assuming the fundamental supermode of the six-core MCF is launched into this taper, the subsequent field evolution is illustrated in Fig. 4.13a–d.

Observing these figures, it becomes clear that the supermode is effectively decomposed into the individual modes associated with each core, enabling their detection by a receiver array positioned at the output. By processing the detected output vector through IDFT, the final result accurately recovers the original input vector as $[1 \ 0 \ 0 \ 0 \ 0 \ 0]$, thus confirming the success of the de-multiplexing technique.

4.2.8 Summary

In conclusion, a rigorous non-orthogonal CMT has been proposed to solve the inaccuracy problem for the conventional CMT under strong coupling between the cores. Analytical formulas are provided to evaluate the propagation constant and field distribution of the super-modes in such MCFs. In addition to the super-mode analysis, a signal precoding technique is proposed which enables simple generation and detection of the supermodes inside the MCF without requirement of the optical components. All of the concepts have been verified by BPM numerical simulations.

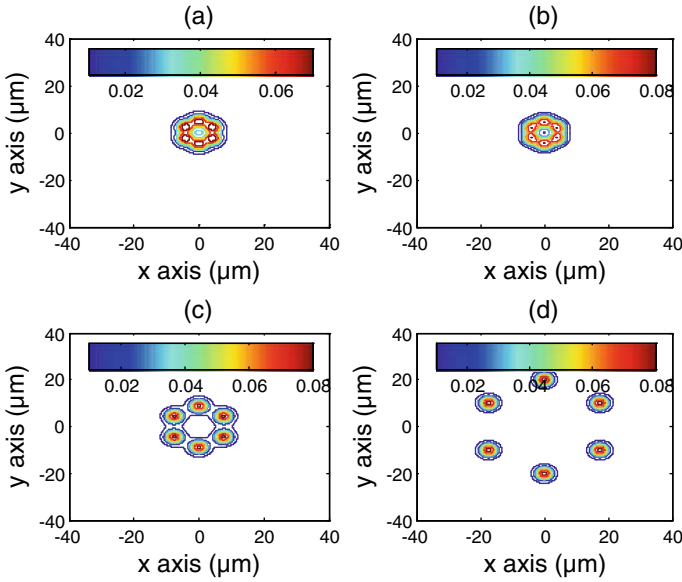


Fig. 4.13 The evolution of the fundamental super-modes of the MCF along the taper **a** at the input of the taper, **b** after propagating in the taper for 100 μm , **c** after in the taper for 2500 μm , **d** after propagating in the taper for 10000 μm . This figure is reprinted/adapted with permission from [Opt. Express 22, 10,815–10,824 (2014)] © The Optical Society of America [30]

4.3 Analytical Study of Super-Modes in Multi-core Fibers with Multimode Cores

The previous two sections deal with the multi-core fibers with the individual single mode cores. In order to increase the system capacity, multimode cores could be used to form the multi-core fibers. There have been few publications on this topic. In this section, a non-orthogonal coupled mode theory is used to derive the rigorous solutions for the super-modes inside multimode multi-core fibers (MMMCF). Analytical formulas are provided for both the propagation constants and the Eigen vectors for the MMMCF super-modes. Numerical simulation verifies the theory with good agreement.

4.3.1 Introduction

There have been significant attentions on the super-modes of the MCFs among the research communities [2, 7, 11–13, 15–21, 23, 24, 29]. Most of them used either numerical [2, 11, 20, 29] or analytical methods [7, 15–19, 21, 23, 24] to characterize the supermodes inside MCFs with single mode cores. Recently, it was proposed

that MCFs with identical multimode cores can also be used for signal transmission and detailed numerical simulation results were provided [12, 13]. By using the few mode fibers (FMFs) as the cores of the MCFs, the number of the supermodes and the transmission capacity can be significantly increased [12, 13].

Although pioneering investigations have been carried out in [12, 13], there are still a few unsolved questions for the MCFs with multimode cores. For instance, what is the relationship between the supermodes of the MMMCFs and their single mode counterpart? Are there any connections between the supermodes inside these two types of the MCFs? Are there any analytical expressions to calculate the propagation constants and the Eigen mode vectors of supermodes in the MCFs with multimode cores?

It is widely known that mode coupling takes place between modes with identical/similar propagation constants [12, 13, 26]. When the difference of the propagation constants is large, the coupling will be averaged to zero during the propagation. Hence, the coupling in MMMCFs should take place between the modes with identical/similar propagation constants in the same core or different cores [12, 13, 17].

In this section, we derive a few analytical expressions to characterize the supermodes inside the MCFs with circularly aligned multimode cores. A rigorous theory based on the non-orthogonal coupled mode theory (NOCMT) [33, 35] is derived along with explicit expressions for the propagation constants and the Eigen vectors of the super-modes. Numerical simulations are provided to verify the theory. Although it is not mentioned here, the results can be easily extended to the case of MCFs with linearly aligned multimode cores.

The section is organized as follows: Sect. 2 derives the basic equations according to the NOCMT; Sects. 3–6 provide the theoretical analysis for the supermodes formed by non-degenerated modes (modes with identical/similar propagation constants are considered as degenerated modes), two degenerated modes, three degenerated modes and L degenerated modes; Sect. 7 presents the numerical simulations to verify the theory provided in the previous sections; and Sect. 8 is the conclusion.

4.3.2 Rigorous Non-Orthogonal Coupled Mode Equations for Super-Modes Characterization in Circular MCFs

Both the vector and the scalar approaches can be used to analyze the supermodes inside the MMMCFs. In comparison with the vector theory, the scalar wave equation is easier to solve and it retains most of the information as well as the accuracy required for the fiber mode analysis. It is widely adopted in the conventional fiber mode analysis and the MCF supermode analysis [2, 7, 11–13, 15–19, 21, 23, 24]. Therefore, it is also used in this work. It should be noted that the results in this work can be easily extended to the case of vector wave equations.

The procedures to derive the non-orthogonal coupled mode equations are similar to those to derive the non-orthogonal coupled mode equations for the single mode MCFs [35], but substantial difference exists since the derivation here abandons the paraxial approximation in [35].

The wave function of the supermodes inside the MCF should satisfy the scalar Holmholtz equation [26, 29, 35]

$$\nabla^2 \varphi + k^2 \varphi = 0 \quad (4.62)$$

It should be noted that they are composed of the modes of the individual cores with identical propagation constant β [26, 29, 35]

$$\varphi = \sum_{n=1}^N c_n \varphi_n e^{i\beta z} \quad (4.63)$$

where c_n stands for the coefficient of the modes of the individual cores, φ_n the mode profile in the cross-section of each core and they satisfy the following equation [26, 29, 35]

$$\nabla_{\perp}^2 \varphi_n + (k_n^2 - \beta^2) \varphi_n = 0 \quad (4.64)$$

where k_n is the free space wave-number multiplied with the local refractive indexes of each individual core and the cladding, $\nabla_{\perp}^2$ the two dimensional Laplace operator.

Substituting Eqs. (4.43–4.64) into (4.62), the following equation can be derived

$$\sum_{n=1}^N \frac{\partial^2 c_n}{\partial z^2} \varphi_n + \sum_{n=1}^N 2j\beta \frac{\partial c_n}{\partial z} \varphi_n = - \sum_{n=1}^N (k^2 - k_n^2) c_n \varphi_n \quad (4.65)$$

Unlike the published NOCMT [29, 33, 35], the paraxial approximation is not necessary to solve the above equation in this rigorous NOCMT. Equation (4.65) can be rewritten into the following matrix form by multiplying the mode profile φ_m^* on both sides of the equation and integrating over the x - y plane.

$$\mathbf{S} \frac{\partial^2 \mathbf{c}}{\partial z^2} + j2\beta \mathbf{S} \frac{\partial \mathbf{c}}{\partial z} = \mathbf{E} \mathbf{c} \quad (4.66)$$

where vector $\mathbf{c}$ has its i th element as c_i and the two matrices $\mathbf{E}$ and $\mathbf{S}$ have the elements as

$$\begin{aligned}
E_{mn} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (k^2 - k_n^2) \varphi_m^* \varphi_n dx dy \\
S_{mn} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \varphi_m^* \varphi_n dx dy
\end{aligned} \tag{4.67}$$

When the paraxial approximation is taken, Eq. (4.66) is reduced to the conventional non-orthogonal coupled mode equation in [29, 33, 35].

Without the paraxial approximation, Eq. (4.66) can be solved analytical by multiplying $\mathbf{S}^{-1}$ on both sides of the equation and perform the matrix decomposition

$$\begin{aligned}
\frac{\partial^2 \mathbf{d}}{\partial z^2} + j2\beta \frac{\partial \mathbf{d}}{\partial z} &= \mathbf{D} \mathbf{d} \\
\mathbf{d} &= \mathbf{T}^{-1} \mathbf{c} \\
\mathbf{S}^{-1} \mathbf{E} &= \mathbf{T} \mathbf{D} \mathbf{T}^{-1}
\end{aligned} \tag{4.68}$$

Since.

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\varphi|^2 dx dy = \mathbf{c}^H \mathbf{S} \mathbf{c} > 0 \tag{4.69}$$

$\mathbf{S}$ is a positive definite matrix, so is $\mathbf{S}^{-1}$.

According to the Green's theorem, we have

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\nabla_{\perp}^2 \varphi_n \varphi_m^* - \nabla_{\perp}^2 \varphi_m^* \varphi_n) dx dy = 0 \tag{4.70}$$

Combining Eq. (4.70) with Eqs. (4.64) and (4.67), it can be inferred that $\mathbf{E}$ is Hermitian. According to the matrix theory, it can be proved that the product of a positive definite matrix and a Hermitian matrix is a matrix with real Eigen values. Hence $\mathbf{S}^{-1} \mathbf{E}$ has real Eigen values as the diagonal elements of matrix $\mathbf{D}$.

Therefore Eq. (4.68) can be solved as

$$\begin{aligned}
\mathbf{d}(L) &= \exp(j\beta^+ z) \mathbf{d}^+(0) + \exp(j\beta^- z) \mathbf{d}^-(0) \\
\beta^{\pm} &= -j\beta \mathbf{I} \pm j\sqrt{\beta^2 \mathbf{I} + 4\mathbf{D}}
\end{aligned} \tag{4.71}$$

Since there are no counter-propagating modes, we have

$$\mathbf{d}(L) = \exp(j\beta^+ L) \mathbf{d}^+(0) \tag{4.72}$$

Hence the amplitude vector $\mathbf{c}$ can be calculated by

$$\mathbf{c}(L) = \mathbf{T} \exp(j\boldsymbol{\beta}^+ L) \mathbf{T}^{-1} \mathbf{c}^+(0) \quad (4.73a)$$

while the propagation constants of the super-modes are

$$\sqrt{\beta^2 + 4D_{nn}} \quad (4.73b)$$

The Eigen vectors for the super-modes should be the row vectors of matrix $\mathbf{T}$. Since $\mathbf{S}^{-1}\mathbf{E}$ is not necessarily Hermitian, the Eigen vectors are not necessarily orthogonal to each other.

4.3.3 Super-Modes Formed by Non-Degenerated Modes

Assuming that the MMMCFs are composed of the equally spaced and circularly aligned multimode fibers, mode coupling will take place between the modes with identical/similar propagation constants. If the modes are non-degenerated, the matrix $\mathbf{E}$ and $\mathbf{S}$ in Eq. (4.68) should be circulant matrices. Therefore, the N elements in one row of the matrices $\mathbf{E}$ and $\mathbf{S}$ can be denoted as $E_1 \dots E_N$, $S_1 \dots S_N$ [35], and they can be decomposed into [34, 35]

$$\begin{aligned} \mathbf{E} &= \mathbf{Q} \mathbf{D}_E \mathbf{Q}^H \\ \mathbf{S} &= \mathbf{Q} \mathbf{D}_S \mathbf{Q}^H \end{aligned} \quad (4.74)$$

where the diagonal matrices $\mathbf{D}_E$ and $\mathbf{D}_S$ have the diagonal elements as [34, 35]

$$\begin{aligned} D_{E,nm} &= \sum_{m=1}^N E_m \exp\left(-j \frac{2\pi(n-1)(m-1)}{N}\right) \\ D_{S,nm} &= \sum_{m=1}^N S_m \exp\left(-j \frac{2\pi(n-1)(m-1)}{N}\right) \end{aligned} \quad (4.75)$$

and the matrix $\mathbf{Q}$ is the normalized IDFT matrix with the elements of [34, 35]

$$Q_{mn} = \frac{1}{\sqrt{N}} \exp\left(j \frac{2\pi(m-1)(n-1)}{N}\right) \quad (4.76)$$

It should be noted that the matrix $\mathbf{Q}$ is neither related to the coupling coefficient nor the overlap integral of the modes. The Eigen vectors are the row vectors of $\mathbf{Q}$. Since there might be degenerated super-modes [35], $\mathbf{Q}$ is not unique, i. e. the rows composed of the degenerated super-modes Eigen vectors can be superposed to get a different $\mathbf{Q}$.

The propagation constants can be solved as

$$\beta_n = \sqrt{\beta^2 + 4 \frac{D_{E,nm}}{D_{S,nm}}} \quad (4.77)$$

If the paraxial approximation is adopted, Eq. (4.77) will be reduced to the formulas derived for the MCFs with single mode cores [35].

4.3.4 Super-Modes Formed by Two Degenerated Modes

If the supermodes are formed by two degenerated modes within the individual cores, the matrices $\mathbf{E}$ and $\mathbf{S}$ should be

$$\begin{aligned} \mathbf{E} &= \begin{pmatrix} \mathbf{E}_{11} & \mathbf{E}_{12} \\ \mathbf{E}_{21} & \mathbf{E}_{22} \end{pmatrix} \\ \mathbf{S} &= \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \end{aligned} \quad (4.78)$$

where $\mathbf{E}_{pq}/\mathbf{S}_{pq}$ stands for the coupling/mode overlapping between the degenerated mode p and mode q among the cores. Since they are circulant matrices, they can be decomposed as

$$\begin{aligned} \mathbf{E}_{pq} &= \mathbf{Q} \mathbf{D}_{E,pq} \mathbf{Q}^H \\ \mathbf{S}_{pq} &= \mathbf{Q} \mathbf{D}_{S,pq} \mathbf{Q}^H \end{aligned} \quad (4.79)$$

As discussed in the previous section, the matrix $\mathbf{Q}$ is not related to the coupling/overlap coefficient. Hence the total coupling/mode overlapping matrix $\mathbf{E}/\mathbf{S}$ can be formulated as

$$\begin{aligned} \mathbf{E}/\mathbf{S} &= \mathbf{Q}_{total} \begin{pmatrix} \mathbf{D}_{E/S,11} & \mathbf{D}_{E/S,12} \\ \mathbf{D}_{E/S,21} & \mathbf{D}_{E/S,22} \end{pmatrix} \mathbf{Q}_{total}^H \\ \mathbf{Q}_{total} &= \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix} \end{aligned} \quad (4.80)$$

Therefore, the matrix $\mathbf{S}^{-1}\mathbf{E}$ can be decomposed into

$$\mathbf{S}^{-1}\mathbf{E} = \mathbf{Q}_{total} \begin{pmatrix} \mathbf{D}_{S,11} & \mathbf{D}_{S,12} \\ \mathbf{D}_{S,21} & \mathbf{D}_{S,22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{D}_{E,11} & \mathbf{D}_{E,12} \\ \mathbf{D}_{E,21} & \mathbf{D}_{E,22} \end{pmatrix} \mathbf{Q}_{total}^H \quad (4.81)$$

It should be noted that since $\mathbf{Q}$ is not unique, the two $\mathbf{Q}$ s in the matrix $\mathbf{Q}_{total}$ can either take the same form or a different form. Through row and column exchange,

the matrix in the middle of Eq. (4.67) can be decomposed into

$$\begin{pmatrix} \mathbf{D}_{S,11} & \mathbf{D}_{S,12} \\ \mathbf{D}_{S,21} & \mathbf{D}_{S,22} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{D}_{E,11} & \mathbf{D}_{E,12} \\ \mathbf{D}_{E,21} & \mathbf{D}_{E,22} \end{pmatrix} = \mathbf{U}\mathbf{N}\mathbf{U}^H \quad (4.82)$$

where $\mathbf{U}$ is a matrix resulting from the row exchange of the identity matrix and $\mathbf{N}$ is a block diagonal matrix as

$$\mathbf{N} = \begin{pmatrix} n_{11}^1 & n_{12}^1 & & \\ n_{21}^1 & n_{22}^1 & & \\ & & \ddots & \\ & & & n_{11}^N & n_{12}^N \\ & & & n_{21}^N & n_{22}^N \end{pmatrix} \quad (4.83)$$

where the block matrices are obtained by

$$\begin{aligned} \begin{pmatrix} n_{11}^l & n_{12}^l \\ n_{21}^l & n_{22}^l \end{pmatrix} &= \begin{pmatrix} d_{S,11}^l & d_{S,12}^l \\ d_{S,21}^l & d_{S,22}^l \end{pmatrix}^{-1} \begin{pmatrix} d_{E,11}^l & d_{E,12}^l \\ d_{E,21}^l & d_{E,22}^l \end{pmatrix} \\ d_{E,pq}^l &= \sum_{m=1}^N E_{m,pq} \exp\left(-j \frac{2\pi(l-1)(m-1)}{N}\right) \\ d_{S,pq}^l &= \sum_{m=1}^N S_{m,pq} \exp\left(-j \frac{2\pi(l-1)(m-1)}{N}\right) \end{aligned} \quad (4.84)$$

where $E_{m,pq}$ and $S_{m,pq}$ are the elements of the circulant matrices $\mathbf{E}$ and $\mathbf{S}$. Matrix $\mathbf{N}$ can be decomposed into

$$\mathbf{N} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} \quad (4.85)$$

where $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are the eigen values of matrix $\mathbf{S}^{-1}\mathbf{E}$, and $\mathbf{V}$ is a block diagonal matrix. With similar mathematical techniques in [21], the elements of $\mathbf{V}$ and $\mathbf{\Lambda}$ can be obtained analytically. Since $\mathbf{N}$ is a block diagonal matrix, its $(2l-1)$ th and $(2l)$ th eigen values (also the $(2l-1)$ th and $(2l)$ th Eigen values of matrix $\mathbf{S}^{-1}\mathbf{E}$) are the roots of the following equation

$$\lambda^2 - (n_{11}^l + n_{22}^l)\lambda + n_{11}^l n_{22}^l - n_{12}^l n_{21}^l = 0 \quad (4.86)$$

and the roots can be evaluated analytically as

$$\lambda_{\pm} = \frac{(n_{11}^l + n_{22}^l) \pm \sqrt{(n_{11}^l - n_{22}^l)^2 + 4n_{12}^l n_{21}^l}}{2} \quad (4.87)$$

the corresponding propagation constants are

$$\sqrt{\beta^2 + 4\lambda_{\pm}} \quad (4.88)$$

while Eigen vectors of the sub-matrix are

$$\left(\frac{(n'_{11} - n'_{22}) \pm \sqrt{(n'_{11} - n'_{22})^2 + 4n'_{12}n'_{21}}}{2n'_{21}}, 1 \right) \quad (4.89)$$

They will form the sub-matrix of the block diagonal matrix $\mathbf{V}$. The total Eigen vectors of the matrix $\mathbf{S}^{-1}\mathbf{E}$ will be the row vectors of matrix $\mathbf{Q}_{\text{total}}\mathbf{UV}$. According to the analysis above, the $(2n-1)$ th and the $(2n)$ th row vectors of matrix $\mathbf{Q}_{\text{total}}\mathbf{UV}$ should be

$$\begin{pmatrix} \mathbf{q}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{q}_n \end{pmatrix} \begin{pmatrix} v_n^{11} & v_n^{12} \\ v_n^{21} & v_n^{22} \end{pmatrix} = \begin{pmatrix} v_n^{11} \mathbf{q}_n & v_n^{12} \mathbf{q}_n \\ v_n^{21} \mathbf{q}_n & v_n^{22} \mathbf{q}_n \end{pmatrix} \quad (4.90)$$

where $\mathbf{q}_n$ is the n th row of $\mathbf{Q}$ indicated by Eq. (4.64) and v_{ij} the elements of the sub-matrix of $\mathbf{V}$ calculated from Eq. (4.89). It can be seen from the above expression that the Eigen vectors of the super-modes generated by the coupling between the two degenerated modes of individual cores are the super-position of the Eigen vectors of the MCFs with single mode cores. It should be noted that $\mathbf{V}$ is not necessarily a unitary matrix, and therefore, the two new Eigen vectors are not necessarily orthogonal to each other. Moreover, since matrix $\mathbf{Q}$ is not unique due to super-modes degeneracy, the two vectors of $\mathbf{q}_n$ in Eq. (4.90) can be either the same or different when $\mathbf{q}_n$ stands for an Eigen vector of a degenerated super-mode of the MCFs with single mode cores. However, the Eigen vectors of the MMMCF are still the super-position of the Eigen vectors of the MCFs with single mode cores.

4.3.5 Super-Modes Formed by Three Degenerated Modes

When the cores have three degenerated modes, the coupling/mode overlapping matrices are

$$\mathbf{E}/\mathbf{S} = \begin{pmatrix} \mathbf{E}/\mathbf{S}_{11} & \mathbf{E}/\mathbf{S}_{12} & \mathbf{E}/\mathbf{S}_{13} \\ \mathbf{E}/\mathbf{S}_{21} & \mathbf{E}/\mathbf{S}_{22} & \mathbf{E}/\mathbf{S}_{23} \\ \mathbf{E}/\mathbf{S}_{31} & \mathbf{E}/\mathbf{S}_{32} & \mathbf{E}/\mathbf{S}_{33} \end{pmatrix} \quad (4.91)$$

Similar to the two-mode case, the matrix $\mathbf{S}^{-1}\mathbf{E}$ can be decomposed as

$$\mathbf{S}^{-1}\mathbf{E} = \begin{pmatrix} \mathbf{Q} & & \\ & \mathbf{Q} & \\ & & \mathbf{Q} \end{pmatrix} \begin{pmatrix} \mathbf{D}_{S,11} & \mathbf{D}_{S,12} & \mathbf{D}_{S,13} \\ \mathbf{D}_{S,21} & \mathbf{D}_{S,22} & \mathbf{D}_{S,23} \\ \mathbf{D}_{S,31} & \mathbf{D}_{S,32} & \mathbf{D}_{S,33} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{D}_{E,11} & \mathbf{D}_{E,12} & \mathbf{D}_{E,13} \\ \mathbf{D}_{E,21} & \mathbf{D}_{E,22} & \mathbf{D}_{E,23} \\ \mathbf{D}_{E,31} & \mathbf{D}_{E,32} & \mathbf{D}_{E,33} \end{pmatrix} \begin{pmatrix} \mathbf{Q}^H & & \\ & \mathbf{Q}^H & \\ & & \mathbf{Q}^H \end{pmatrix} \quad (4.92)$$

Similarly, the three $\mathbf{Q}$ s in the expression can either take the same form or a different form. The matrix in the middle can be decomposed into

$$\begin{pmatrix} \mathbf{D}_{S,11} & \mathbf{D}_{S,12} & \mathbf{D}_{S,13} \\ \mathbf{D}_{S,21} & \mathbf{D}_{S,22} & \mathbf{D}_{S,23} \\ \mathbf{D}_{S,31} & \mathbf{D}_{S,32} & \mathbf{D}_{S,33} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{D}_{E,11} & \mathbf{D}_{E,12} & \mathbf{D}_{E,13} \\ \mathbf{D}_{E,21} & \mathbf{D}_{E,22} & \mathbf{D}_{E,23} \\ \mathbf{D}_{E,31} & \mathbf{D}_{E,32} & \mathbf{D}_{E,33} \end{pmatrix} = \mathbf{U}\mathbf{N}\mathbf{U}^H \quad (4.93)$$

where the block diagonal matrix $\mathbf{N}$ is

$$\mathbf{N} = \begin{pmatrix} n_{11}^1 & n_{12}^1 & n_{13}^1 & & & \\ n_{21}^1 & n_{22}^1 & n_{23}^1 & & & \\ n_{31}^1 & n_{32}^1 & n_{33}^1 & & & \\ & & & \ddots & & \\ & & & & n_{11}^N & n_{12}^N & n_{13}^N \\ & & & & n_{21}^N & n_{22}^N & n_{23}^N \\ & & & & n_{31}^N & n_{32}^N & n_{33}^N \end{pmatrix} \quad (4.94)$$

where the sub-matrices inside $\mathbf{N}$ are obtained by

$$\begin{pmatrix} n_{11}^l & n_{12}^l & n_{13}^l \\ n_{21}^l & n_{22}^l & n_{23}^l \\ n_{31}^l & n_{32}^l & n_{33}^l \end{pmatrix} = \begin{pmatrix} d_{S,11}^l & d_{S,12}^l & d_{S,13}^l \\ d_{S,21}^l & d_{S,22}^l & d_{S,23}^l \\ d_{S,31}^l & d_{S,32}^l & d_{S,33}^l \end{pmatrix}^{-1} \begin{pmatrix} d_{E,11}^l & d_{E,12}^l & d_{E,13}^l \\ d_{E,21}^l & d_{E,22}^l & d_{E,23}^l \\ d_{E,31}^l & d_{E,32}^l & d_{E,33}^l \end{pmatrix} \quad (4.95)$$

$$d_{E,pq}^l = \sum_{m=1}^N E_{m,pq} \exp\left(-j \frac{2\pi(l-1)(m-1)}{N}\right)$$

$$d_{S,pq}^l = \sum_{m=1}^N S_{m,pq} \exp\left(-j \frac{2\pi(l-1)(m-1)}{N}\right)$$

The Eigen values of matrix $\mathbf{N}$ can be obtained by calculating the Eigen values of the sub-matrix

$$\begin{pmatrix} n_{11}^l & n_{12}^l & n_{13}^l \\ n_{21}^l & n_{22}^l & n_{23}^l \\ n_{31}^l & n_{32}^l & n_{33}^l \end{pmatrix} \quad (4.96)$$

As indicated in [15], the Eigen values of the above matrix can be obtained by calculating the roots of the characteristic equation

$$\begin{aligned}
 \lambda^3 + A\lambda^2 + B\lambda + C &= 0 \\
 A &= -(n_{11}^l + n_{22}^l + n_{33}^l) \\
 B &= (n_{11}^l n_{22}^l + n_{11}^l n_{33}^l + n_{22}^l n_{33}^l - n_{12}^l n_{21}^l - n_{23}^l n_{32}^l - n_{13}^l n_{31}^l) \\
 C &= -n_{11}^l n_{22}^l n_{33}^l + n_{11}^l n_{23}^l n_{32}^l + n_{22}^l n_{13}^l n_{31}^l + n_{33}^l n_{12}^l n_{21}^l - n_{12}^l n_{23}^l n_{31}^l - n_{13}^l n_{32}^l n_{21}^l
 \end{aligned} \tag{4.97}$$

The roots of the equation can be solved analytically as [15, 24]

$$\begin{aligned}
 \lambda_k &= -\frac{1}{3} \left(b + u_k D + \frac{\Delta_0}{u_k D} \right) \\
 u_1 &= 1, u_2 = \frac{-1 + j\sqrt{3}}{2}, u_3 = \frac{-1 - j\sqrt{3}}{2} \\
 D &= \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}} \\
 \Delta_0 &= B^2 - 3C \\
 \Delta_1 &= 2B^3 - 9BC + 27d
 \end{aligned} \tag{4.98}$$

The propagation constants can be obtained by expressions similar to Eq. (4.88).

With those Eigen values, one is able to obtain the corresponding Eigen vectors of the matrix in Eq. (4.96) and they will form the sub-matrix of $\mathbf{V}$. The total Eigen vectors of $\mathbf{S}^{-1}\mathbf{E}$ will still be formed by the superposition of the Eigen vectors of $\mathbf{Q}$ with the coefficients from the sub-matrix of $\mathbf{V}$, as indicated by the following equation

$$\begin{pmatrix} \mathbf{q}_n & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{q}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{q}_n \end{pmatrix} \begin{pmatrix} v_n^{11} & v_n^{12} & v_n^{13} \\ v_n^{21} & v_n^{22} & v_n^{23} \\ v_n^{31} & v_n^{32} & v_n^{33} \end{pmatrix} = \begin{pmatrix} v_n^{11} \mathbf{q}_n & v_n^{12} \mathbf{q}_n & v_n^{13} \mathbf{q}_n \\ v_n^{21} \mathbf{q}_n & v_n^{22} \mathbf{q}_n & v_n^{23} \mathbf{q}_n \\ v_n^{31} \mathbf{q}_n & v_n^{32} \mathbf{q}_n & v_n^{33} \mathbf{q}_n \end{pmatrix} \tag{4.99}$$

As indicated before, when the supermodes of the single mode MCF are degenerated, the Eigen vectors $\mathbf{q}_n$ are not unique, and they can be either the same or different in Eq. (4.99).

4.3.6 Super-Modes Formed by L Degenerated Modes

When the cores have L degenerated modes, the matrices $\mathbf{E}$ and $\mathbf{S}$ will be formed by L^2 submatrices

$$\mathbf{E}/\mathbf{S} = \begin{pmatrix} \mathbf{E}/\mathbf{S}_{11} & \mathbf{E}/\mathbf{S}_{12} & \cdots & \mathbf{E}/\mathbf{S}_{1L} \\ \mathbf{E}/\mathbf{S}_{21} & \mathbf{E}/\mathbf{S}_{22} & \cdots & \vdots \\ \vdots & & \ddots & \mathbf{E}/\mathbf{S}_{L-1L} \\ \mathbf{E}/\mathbf{S}_{L1} & \cdots & \mathbf{E}/\mathbf{S}_{LL-1} & \mathbf{E}/\mathbf{S}_{LL} \end{pmatrix} \quad (4.100)$$

Similarly, the matrix $\mathbf{S}^{-1}\mathbf{E}$ can be related to a block diagonal matrix $\mathbf{N}$ by

$$\mathbf{S}^{-1}\mathbf{E} = \mathbf{Q}_{Total}\mathbf{U}\mathbf{N}\mathbf{U}^H\mathbf{Q}_{Total}^H \quad (4.101)$$

where the block diagonal matrix $\mathbf{N}$ is formed by the following LXL sub-matrices

$$\begin{pmatrix} n_{11}^l & n_{12}^l & \cdots & n_{1L}^l \\ n_{21}^l & n_{22}^l & \cdots & \vdots \\ \vdots & & \ddots & n_{L-1L}^l \\ n_{L1}^l & \cdots & n_{LL-1}^l & n_{LL}^l \end{pmatrix} \quad (4.102)$$

By obtaining the Eigen values of the matrix in Eq. (4.84), one is able to determine the Eigen values of matrix $\mathbf{S}^{-1}\mathbf{E}$. The corresponding propagation constants and the Eigen vectors can also be obtained afterwards. Although there is no direct analytical approach to evaluate the propagation constants and the Eigen vectors, fast numerical techniques are widely used for such problems. It can be inferred from the derivation that the Eigen vectors are still the super-position of the Eigen vectors resulting from the MCFs with single mode cores.

4.3.7 Numerical Simulations and Discussions

A closely arranged six-core MMMCF aligned in a circular array is analyzed as an example. The MMMCF has the core index of 1.45 and the cladding index of 1.4674. These values are from [10] to make the example more realistic and sound. The radii of the cores are 10 μm and the distances between the centers of the cores are also 10 μm . This indicates that there are no gaps between the cores and corresponds to a strong coupling between the cores. The individual cores have multiples modes, such as mode LP01, and LP11a as well as LP11b. Higher order modes are not considered. The cross section of the MCF and the modes of individual cores are plotted in Fig. 4.14.

As discussed in the previous sections, the coupling between the modes with identical/similar propagation constants will form the supermodes inside the MMMCF. The supermodes formed by the coupling between the LP01 modes within the cores corresponds to the analysis in Sect. 4.3, while the supermodes formed by the coupling between modes LP11a and LP11b corresponds to the analysis in Sect. 4.4. To examine

Fig. 4.14 Cross section of the MCF and the mode profiles for each individual cores

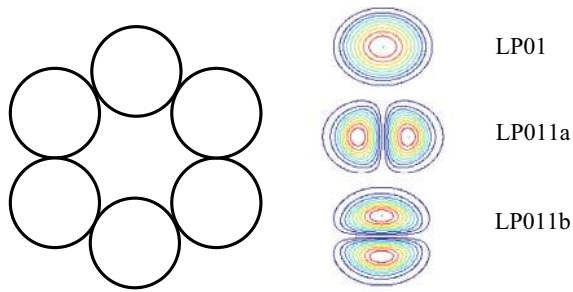


Table 4.7 The effective indexes of the supermodes formed by mode LP01

Mode No	BPM	NOCMT
1	1.4650	1.4649
2	1.4648	1.4646
3	1.4648	1.4646
4	1.4642	1.4640
5	1.4642	1.4640
6	1.4637	1.4637

the formulas derived in these sections, three dimensional beam propagation method (3DBPM) is used to calculate the effective refractive indexes of these supermodes and the results are compared with the ones from the analytical expression in these sections.

The refractive indexes calculated by the 3DBPM and the non-orthogonal CMT (NOCMT) are illustrated in Tables 4.7 and 4.8. The modes are arranged in a descending order according to their effective indexes. It can be observed that the results from the NOCMT have a very close match with the numerical simulations, and therefore the formulas derived in this section will be a powerful tool to analyze the MCFs with multimode cores.

The NOCMT can predict not only the propagation constant of the MMMCF, but also the eigen vectors of the propagation matrix which describe the super-mode amplitudes distribution. These vectors can be obtained by Eqs. (4.76) and (4.90). However, when the super-modes are formed by degenerated modes of individual cores, although they can be the super-position of the corresponding Eigen vectors resulting from the single mode MCF, the coefficients for the superposition might not be unique. Therefore, to obtain a close match with the BPM simulation results, the coefficients for the superposition should be chosen carefully.

The super-modes profiles predicted by the NOCMT are plotted in Figs. 4.15, 4.16, 4.17, 4.18, 4.19, 4.20. Figures 4.15 and 4.16 present the super-modes formed by LP 01, while Figs. 4.17, 4.18, 4.19 and 4.20 present the super-modes formed by LP 11a and LP 11b. The results in Figs. 4.17, 4.18, 4.19 and 4.20 demonstrate the accuracy of the formulas for Eigen vector prediction. It should be noted the matching

Table 4.8 The effective indexes of the supermodes formed by mode LP11a and mode LP11b

Mode No	BPM	NOCMT
7	1.4615	1.4614
8	1.4609	1.4609
9	1.4609	1.4609
10	1.4603	1.4604
11	1.4603	1.4603
12	1.4603	1.4603
13	1.4589	1.4589
14	1.4589	1.4589
15	1.4587	1.4587
16	1.4586	1.4585
17	1.4586	1.4585
18	1.4579	1.4578

between the numerical results and the predicted mode profiles in Figs. 4.17, 4.18, 4.19 and 4.20 might be improved by further adjusting the superposition coefficients.

To quantify the similarity between the modes obtained by the analytical method and the BPM simulations, the correlation coefficient, which is defined as the normalized overlap integrals, is calculated. The correlation coefficients for the modes 1–6 in Figs. 4.14 and 4.15 are 0.99, 0.99, 0.99, 0.98, 0.98 and 0.99; the correlation coefficients for the modes 7–12 in Figs. 4.16 and 4.17 are 0.98, 0.95, 0.95, 0.98, 0.94 and 0.95; the correlation coefficients for the modes 13–18 in Figs. 4.18 and 4.19 are 0.98, 0.98, 0.99, 0.96, 0.97 and 0.99. These values demonstrate the close match between supermodes obtained by the both methods.

4.3.8 Summary

In summary, we have proposed an analytical tool to analyze the supermodes of the MMMCFs. Explicit expressions for the propagation constants and the eigen vectors are derived by analyzing the coupling matrix of the MCFs with circularly aligned multimode cores. Numerical results are compared with the analytical predictions, and the accuracy of the formulas is demonstrated. Those expressions will be extremely helpful for the design and analysis of the MMMCFs.

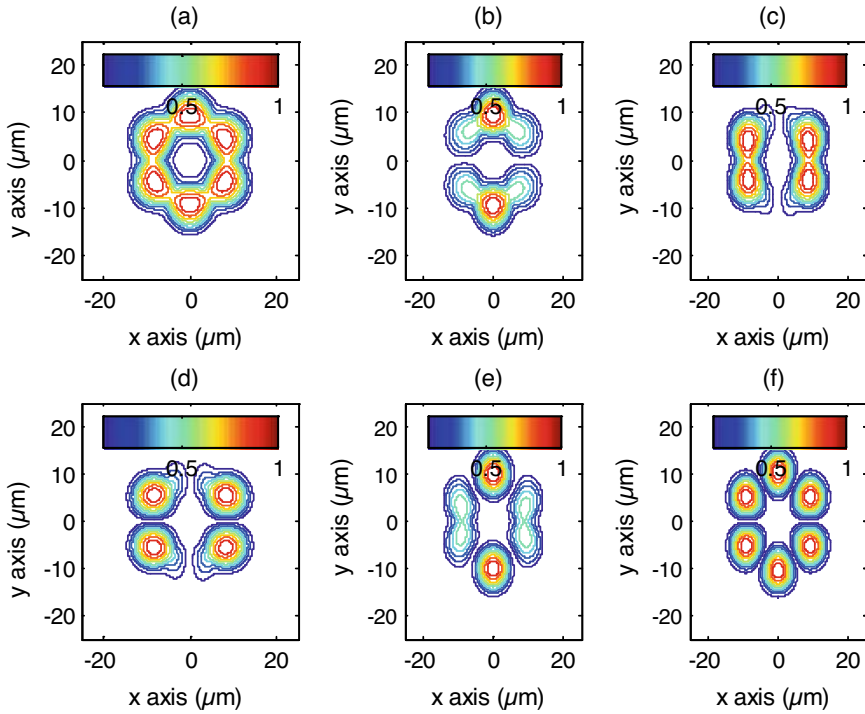


Fig. 4.15 The supermodes formed by mode LP 01 and obtained by BPM (a-f) mode profiles of the supermodes 1–6

4.4 Power Splitting and Switching in a Multi-core Fiber Based on the Multimode Interference Effect

Part of the section reprinted/adapted with permission from [Opt. Express 23, 22,098–22,107 (2015)] © The Optical Society of America [36].

After discussing the super-modes in the multi-core fibers, we move to investigate the functional device based on the structure. In this section, A novel 1 to N optical power splitter and a 1 to N optical switch is proposed based on a multi-core fiber (MCF) featuring a circular cores array, which exploits the multimode interference (MMI) effect in a ring shape fiber connecting the MCF. The MMI effect will convert one image into N output images in the ring and therefore, the ring shape MMI coupler can act as a 1 to N power splitter. One may implement two ring shape MMI couplers along with a phase shifter array between them to realize the optical switch. The output port number can be tuned by altering the phase in the middle of the device.

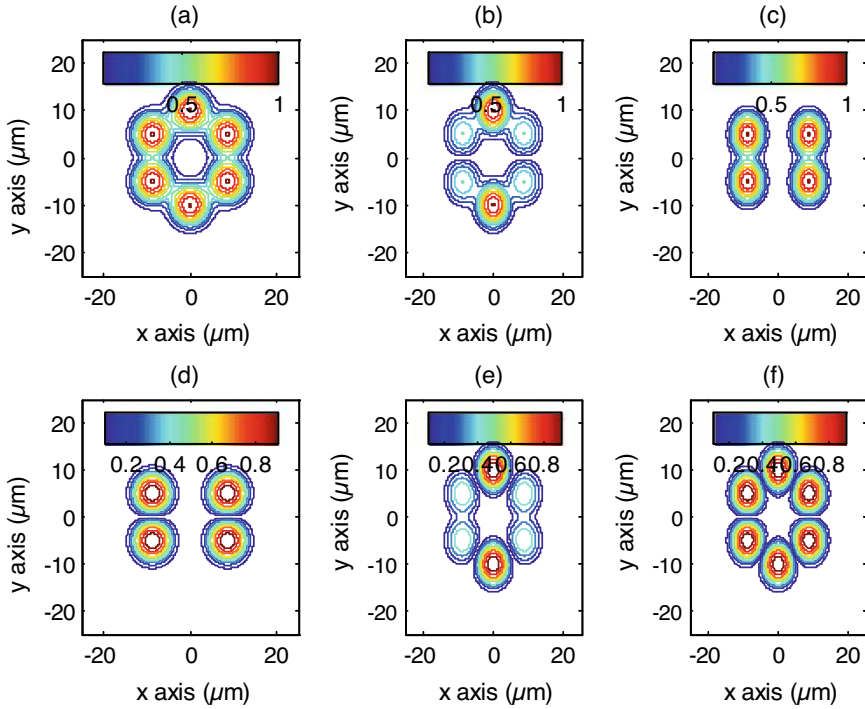


Fig. 4.16 The supermodes formed by mode LP 01 and obtained by Eq. (4.76), **a–f** mode profiles of the supermodes 1–6

4.4.1 Introduction

The MCF cores can have different arrangement, such as the hexagon lattice [5, 6], the linear array [7] or the circular array [8]. While the ring shape core arrangement has been applied in signal transmission, it might be more compatible with the innovative optical network architecture as is discussed below.

The modern optical network features reconfigurability, which is made possible by the optical add/drop multiplexers (ROADMs) [37]. ROADMs are composed of devices called multi-cast switches made of 1 to N power splitters and 1 to N optical switches [37]. Such optical components are usually single mode fiber (SMF) based, which poses a significant challenge in incorporating MCFs into these advanced network structures, as their adaptation becomes quite difficult. To achieve the splitting of signals propagating within a specific core of an MCF into multiple other cores, or to switch the signals between cores within the MCF, a substantial number of MCF-to-SMF couplers are needed.

By implementing power splitting and optical switching with a compact structure that seamlessly integrates with MCF technology, significant efforts in incorporating

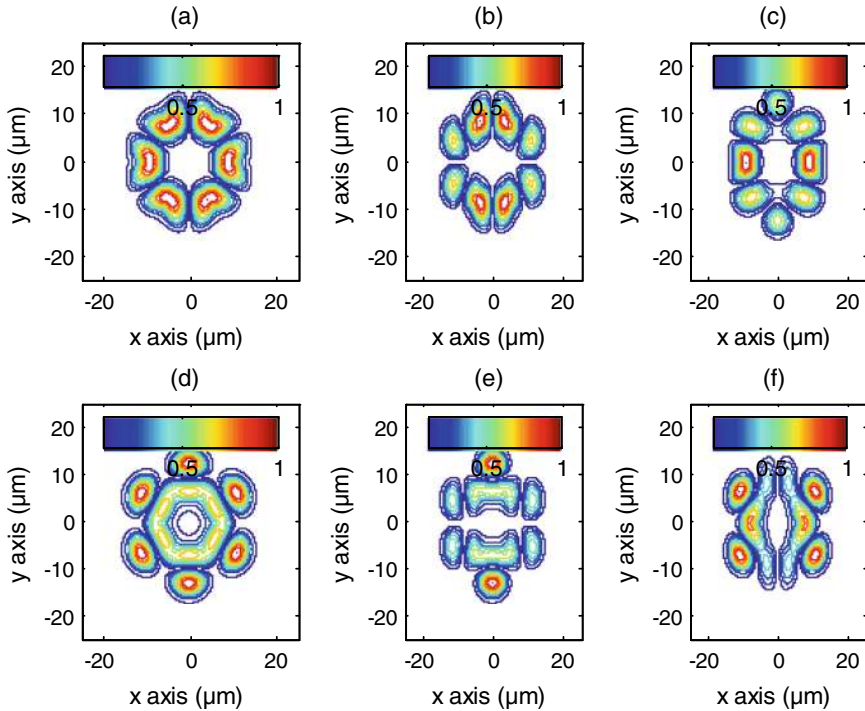


Fig. 4.17 The supermodes formed by mode LP 11a and mode LP 11b and obtained by BPM, **a–f** mode profiles of the supermodes 7–12

MCFs into innovative network structures can be mitigated. In this section, we introduce a novel approach to realize power splitters and optical switches specifically tailored for MCFs based on the MMI effect in a ring core fiber.

The MMI effect, which is well-established in planar waveguides, possesses the capability to transform a single input image into multiple replicated images, each with distinct phases. This remarkable phenomenon has inspired the development of numerous functional devices, including optical power splitters [38], optical switches [39], and mode converters [40].

Recently, Zhou et al. have found that MMI effect does occur in ring core fibers [41]. Analogous to the planar waveguides, the MMI effect within a ring core fiber also converts a solitary input image into multiple replicated images. Hence, it is natural to be inferred from the planar waveguide case that ring core fiber can be used for power splitters and optical switches. Such optical components are compatible with the MCFs and may be used to as RODAMs for MCF based optical networks.

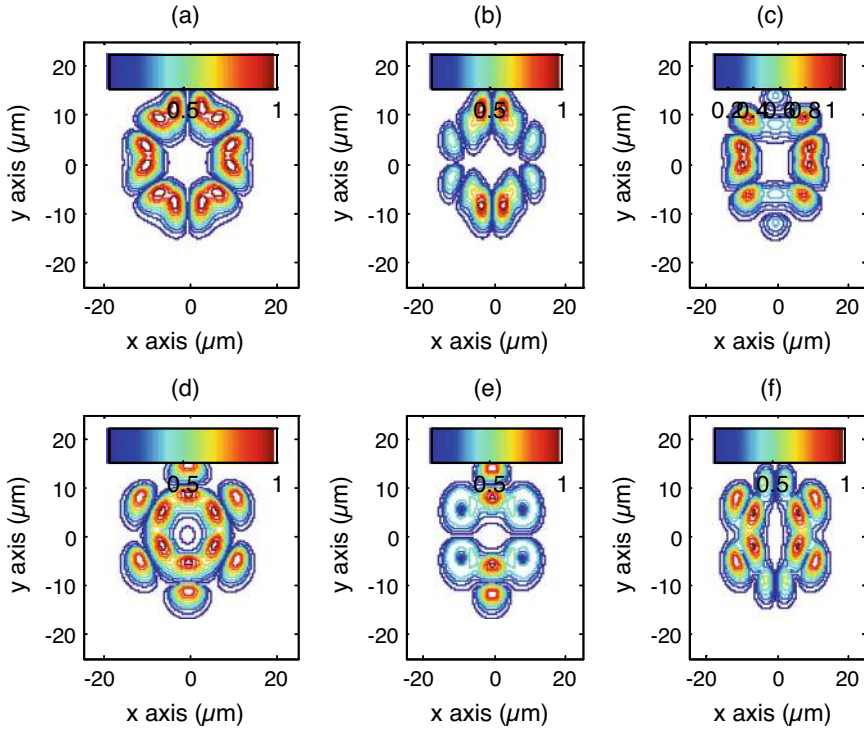


Fig. 4.18 The supermodes formed by mode LP 11a and mode LP 11b and obtained by Eq. (4.90), **a–f** mode profiles of the supermodes 7–12

4.4.2 Theory for the MMI Effect Inside a Ring Core Fiber

A comprehensive explanation of the theory governing the MMI process within ring core fibers is presented in [41]. To facilitate understanding for readers, a concise overview is provided here. When the ring structure is considered “thin,” meaning the radius of the ring (denoted as a) significantly exceeds its width, and the ring core fiber exhibits single-mode behavior radially while being multimode in the angular domain, the MMI effect emerges in case a specific length of the ring-shaped MMI coupler [41] is given:

$$L_N = \frac{2ka^2\pi}{N} \quad (4.103)$$

where N is the number of duplicated images to be generated, $k = nk_0$, n the refractive index of the ring core, k_0 the free space wave number, a the radius of the ring.

Denoting the coefficients for fields at input/output ports as vector $\mathbf{a/b}$ with N elements, the input and the output vectors can be associated by the transfer matrix $\mathbf{T}$

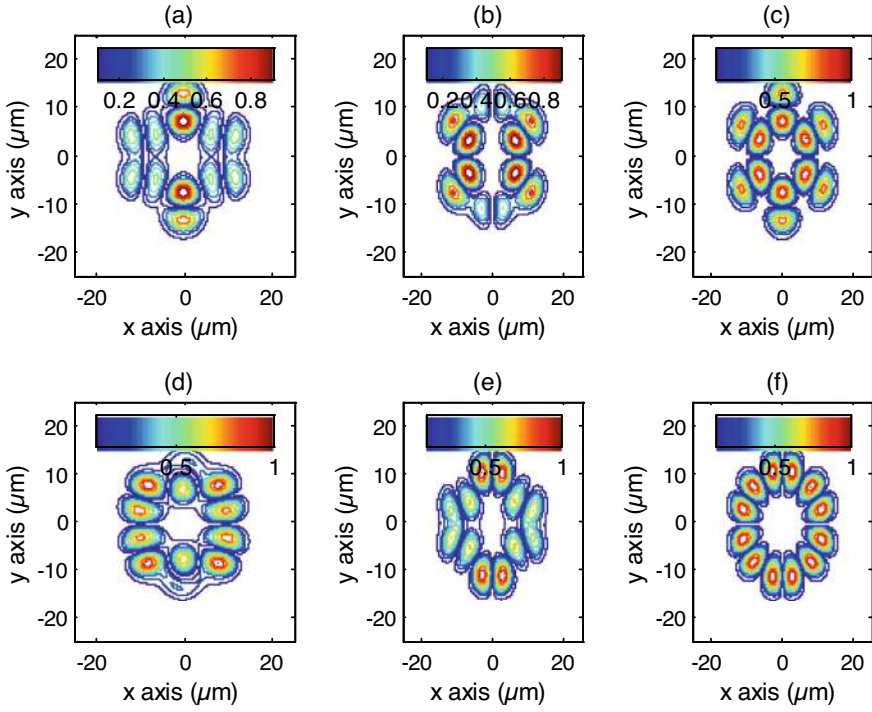


Fig. 4.19 The supermodes formed by mode mode LP 11a and mode LP 11b and obtained by BPM, **a–f** mode profiles of the supermodes 13–18

[41]

$$\mathbf{b} = \mathbf{T}\mathbf{a} \quad (4.104)$$

where matrix $\mathbf{T}$ has been formulated analytically in [41]

$$\mathbf{T}_{mn} = \frac{\exp(-j\frac{\pi}{4})}{\sqrt{N}} \exp\left(j\frac{m^2\pi}{N}\right) \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{n^2\pi}{N}\right) \quad (4.105)$$

$$m = 0, \dots, N-1$$

$$n = 0, \dots, N-1$$

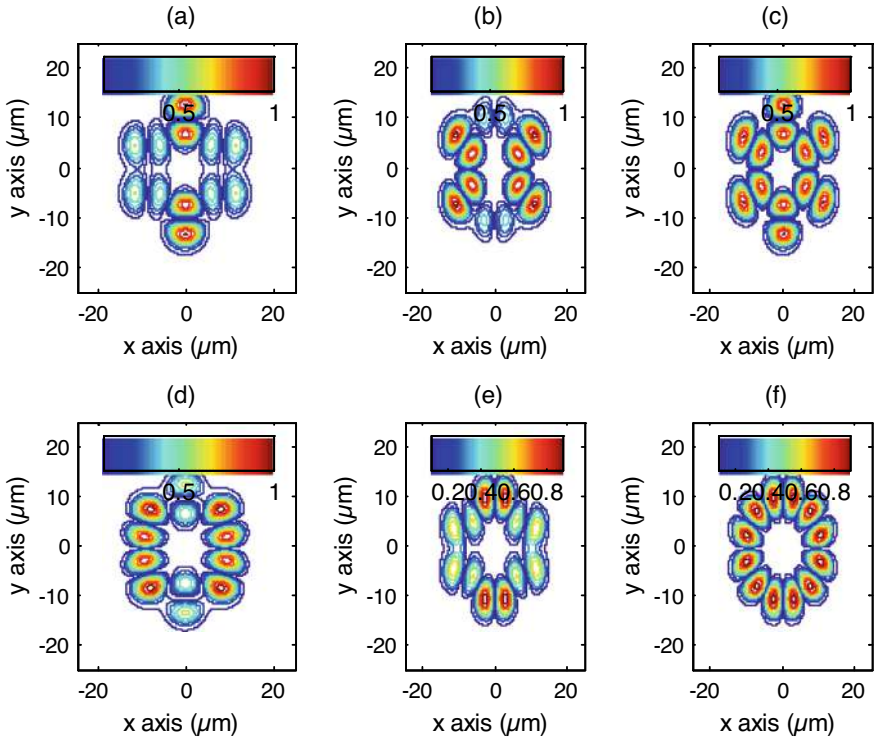


Fig. 4.20 The supermodes formed by mode LP 11a and mode LP 11b and obtained by Eq. (4.90), **a–f** mode profiles of the supermodes 13–18

4.4.3 Working Principle of the Ring Shape MMI Coupler Based Power Splitter and Switch

Based on Eqs. (4.104–4.105), when a signal is injected into one of the input ports of the ring core fiber, the beam will evolve into N replicas at specific length. Obviously, the device is able to convert one input into N outputs, which is a power splitter, specifically designed for a MCF with a circular core array.

An optical switch can be realized with two MMI couplers and phase shifters, just like the planar waveguide case. If the phase shifter array with the following phase shift emerge at the output of the MMI coupler [41]:

$$\exp\left(-j\frac{n^2\pi}{N}\right) \quad (4.106)$$

One will have a modified transfer matrix $\mathbf{T}'$ [41] with the element as [41]:

$$\mathbf{T}'_{mn} = \frac{\exp(-j\frac{\pi}{4})}{\sqrt{N}} \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{m^2\pi}{N}\right) \quad (4.107)$$

If a signal inputs at the m th input of the first ring core fiber, the phase distribution at the output of the ring core fiber obeys $-2\pi mn/N$ by ignoring the constant phase term. From Eq. (4.108), it can be deduced that if N outputs with the phase shift of $-2\pi mn/N$ are injected into the second ring core fiber, the output beam will focus primarily on the $(N-m)$ th output port, suggesting that the m th input port has been switched to the $(N-m)$ th output port.

$$\sum_{n=0}^{N-1} \frac{1}{N} \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(-j\frac{2m'n\pi}{N}\right) = \begin{cases} 1 & m' = N - m \\ 0 & m' \neq N - m \end{cases} \quad (4.108)$$

In order to tune the proposed switch, a tunable phase shifter array is used between the two ring core fibers, which can change the phases as $-2\pi m'n/N$ plus some fixed phase terms, where m' is different integer from m . When the signal appears at the m th input port, the $(N-m')$ th output port the proposed switch will have the corresponding output beam.

The phase shifters have a simple control mechanism, as their tunable part of the phase shift is

$$(0, \varphi_0, \dots, n\varphi_0, \dots, (N-1)\varphi_0) \\ \varphi_0 = \frac{2\pi l}{N} \quad (4.109)$$

where l is an integer, which is the only control parameter. Thermal-optical effect as well as electro-optical effect can be used for phase shifter tuning from $-2\pi mn/N$ to $-2\pi(m-l)n/N$.

4.4.4 Numerical Simulations

Numerical simulations utilizing the Beam Propagation Method (BPM) have been conducted to validate the concept introduced in this study. As an illustrative example, a MCF with eight cores is employed. The cladding of this MCF has a refractive index of 1.45, while the cores exhibit a refractive index difference of 1.2% relative to the cladding [10]. Given the relatively small index difference, the polarization effects are considered negligible. Each core possesses a radius of $2.5 \mu\text{m}$. The eight cores are evenly arranged in a circular layout with a radius of $48 \mu\text{m}$, resulting in a neighboring core spacing, or core pitch, of $36.7 \mu\text{m}$ [8] to achieve satisfactory crosstalk characteristics.

A MMI device is employed utilizing a ring core fiber that possesses identical refractive indices for both its core and cladding regions with respect to the MCF. The

ring is designed with a radius of 48 μm and a width of 5 μm , ensuring compatibility with the dimensions of the MCF. Operating at the signal wavelength of 1550 nm, the length of this ring-shaped MMI coupler is calculated using Eq. (4.103) and subsequently optimized through the simulation, resulting in a final length of 10,858 μm .

The amount of the radial directional modes can be calculated by the normalized frequency V as well as the m th order radial mode cut-off frequency V_m [42]

$$\begin{aligned} V &= kb\sqrt{n_2^2 - n_1^2} \\ V_m &= \frac{(m-1)\pi}{1-\rho} \end{aligned} \quad (4.110)$$

where, b is the outer radius of the ring, ρ the ratio between the inner radius and the outer radius of the ring. In designing the MMI device, Eq. (4.110) serves as a valuable tool to assess its operational viability. Furthermore, a thorough examination of the angular modes' cut-off frequencies is crucial to guarantee the presence of a sufficient number of angular modes, essential for achieving a high-quality MMI process. Based on the author's research, the normalized frequency V must satisfy a specific inequality to optimize the device's performance, ensuring optimal outcomes:

$$\frac{\pi}{1-\rho} < V = kb\sqrt{n_2^2 - n_1^2} < \frac{2\pi}{1-\rho} \quad (4.111)$$

The port labeling principle of the ring core fiber is shown in Fig. 4.21a.

One ring core fiber is enough to form a power splitter, whose schematic is shown in Fig. 4.21b. The beam is injected into one input port, for instance, the 0th port. It evolves in the ring core fiber as shown in Fig. 4.22b, c, which gradually changes to 8 replicas at the output ports in Fig. 4.22d. The distribution of the optical power at the output port is quite even, with the insertion loss of each port very near 9 dB. The deviation in loss for each port remains below 0.3 dB.

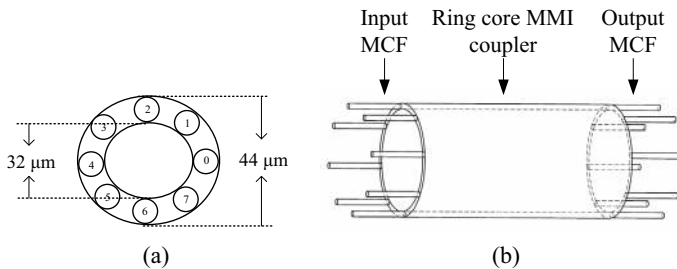


Fig. 4.21 **a** Cross section of the MMI device used for the power splitter and the optical switch **b** schematic of the power splitter. This figure is reprinted/adapted with permission from [Opt. Express 23, 22098-22107 (2015)] © The Optical Society of America [36]

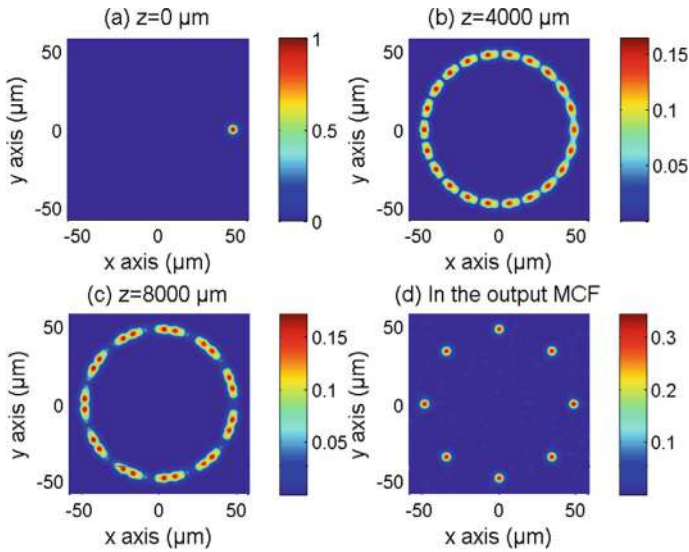


Fig. 4.22 Input and output field of the MMI coupler and the evolution of the field inside the ring core fiber if the input is injected into the 0th input port. This figure is reprinted/adapted with permission from [Opt. Express 23, 22,098–22,107 (2015)] © The Optical Society of America [36]

After the successful demonstration for the MCF power splitter, the optical switch compatible with the MCFs is studied, which consists of four parts. The first ring core fiber transforms one input beam into 8 replicas. The phase of each beam will be changed by the phase shifter arrays. The optical field will concentrate on one of the output ports of the 2nd ring core fiber. Figure 4.23 demonstrates the device structure.

Among the phase shifter arrays, the fixed array should have the following phase shift

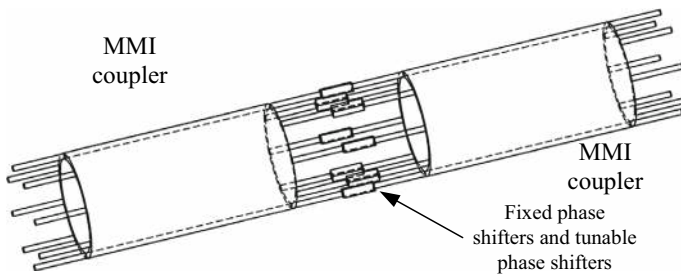


Fig. 4.23 The basic device structure of the proposed optical switch. This figure is reprinted/adapted with permission from [Opt. Express 23, 22,098–22,107 (2015)] © The Optical Society of America [36]

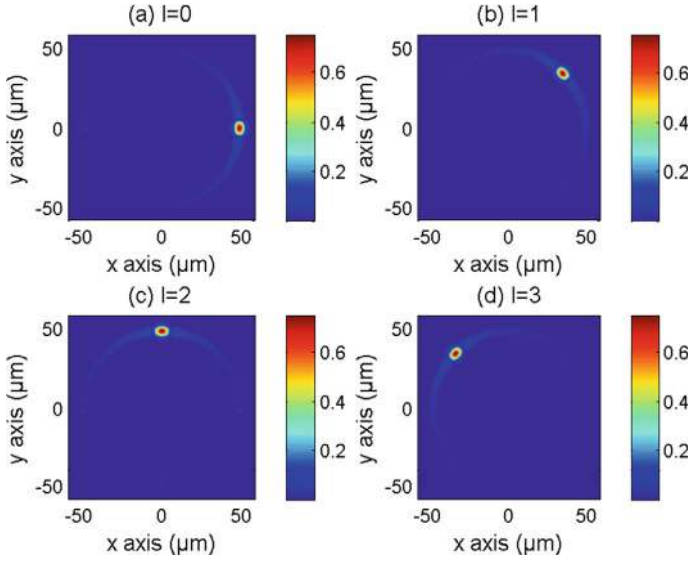


Fig. 4.24 The output of the switch when the input is injected into the 0th input port and the control number l varies from 0 to 3. This figure is reprinted/adapted with permission from [Opt. Express 23, 22,098–22,107 (2015)] © The Optical Society of America [36]

$$\exp\left(-j\frac{2n^2\pi}{N}\right) \quad (4.112)$$

which is twice of the phase shift in Eq. (4.106), and it is the result of adopting two ring shape MMI couplers. The tunable part of the phase shifter array should have the phase shift as

$$\exp\left(j\frac{2ln\pi}{N}\right) \quad (4.113)$$

where l is an integer as the control parameter. When signals are introduced into the m th input port, the corresponding output signal emerges from the $(N-m+l)$ th output port, with a special consideration: if the calculated port number $N-m+l$ is either equal to or greater than N , or less than 0, the actual output port number is determined by taking the remainder of $(N-m+l)$ divided by N . By adjusting the control parameter, denoted as l , it becomes feasible to reroute signals from any core within the MCF to any other core within the same MCF. This mechanism enables the realization of a 1-to- N optical switch.

The implementation of a phase shifter array can be done through diverse methods, including the utilization of external phase shifters [43], arrays of waveguides with varying lengths [11], and waveguides with differentiated refractive indices [10]. In this research, we adopt the approach that combines a MCF featuring cores with distinct refractive indices, followed by an externally tunable phase shifter array. This setup allows for simple adjustment of the control parameter l . During the BPM simulations, the tunable phase shifters are modeled as an array of ideal phase plates, capable of modifying the phase of beams propagating through the waveguides. The array spans $200\text{ }\mu\text{m}$ in length, with the 0th phase shifter possessing the lowest core index of 1.4674 (a 1.2% index difference from a reference), and the 1st phase shifter having an index of 1.4662 (a 1.12% index difference). The remaining phase shifters are assigned corresponding indices accordingly. To mitigate potential reflections arising from the index mismatch between the phase shifters and the ring core fiber, anti-reflection coatings are applied prior to their integration. It is worth noting that the tunable phase shifter array can also be realized through alternative methods.

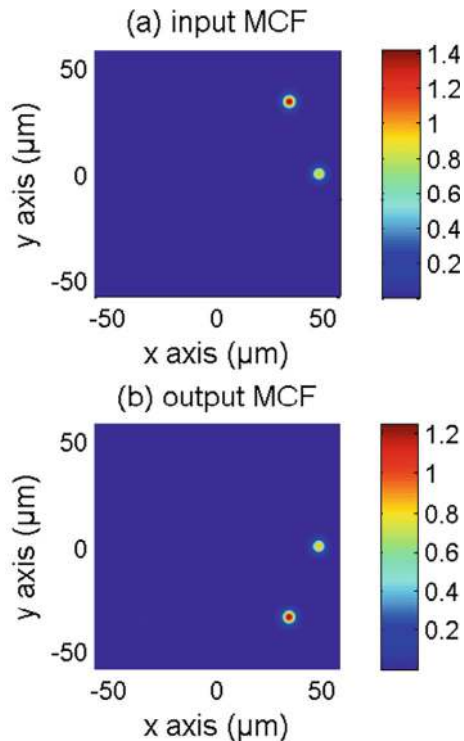
The simulation results demonstrating the operation of the optical switch are shown in Fig. 4. It is assumed that the signal is injected into the 0th input port of the switch, whereas the figures show the output beam emerged from different output ports with the different control parameter l ranging from 0 to 3. The important parameter for the optical switch, i.e., isolation, is quite high. In the most unfavorable scenario, the isolation level attains a value of 28dB.

To show the capabilities of the MCF switch in handling multiple input signals distributed across distinct cores, let's consider a scenario where two input signals are injected into the 0th and 1st input ports of the device, respectively. Figure 4.25a illustrates the input signal distribution, where a power ratio of 1:2 is assumed between the two signals. Setting the control parameter l as 0, the signals are anticipated to emerge from the 0th and (N-1)th, or specifically, the 7th core at the output. The results, as depicted in Fig. 4.25b, are in concordance with the theoretical predictions, validating the functionality of the MCF switch under these conditions.

4.4.5 Summary

In conclusion, we have introduced an innovative 1-to-N power splitter and optical switch specifically designed for circularly aligned MCFs, utilizing the MMI effect within a ring core fiber. The simplicity of these devices is evident; the power splitter

Fig. 4.25 The input and output signal amplitudes in the MCF before and after switching, the control number l is 0. This figure is reprinted/adapted with permission from [Opt. Express 23, 22,098–22,107 (2015)] © The Optical Society of America [36]



is directly implemented using a ring-shaped MMI coupler, while the optical switch comprises two ring-shaped MMI couplers, a fixed phase shifter array, and a tunable phase shifter array. Comprehensive numerical simulations have been conducted to substantiate the functionality of these devices. Furthermore, the concept can be adapted for MCFs with cores arranged in multiple rings, expanding its applicability. The proposed devices hold significant potential in facilitating the integration of MCFs into contemporary optical networks.

References

1. T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, E. Sasaoka, Ultra-low-crosstalk multi-core fiber feasible to ultra-long-haul transmission, in *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference* (2011), pp. 1–3
2. C. Xia, N. Bai, I. Ozdur, X. Zhou, G. Li, Supermodes for optical transmission. *Opt. Express* **19**(17), 16653–16664 (2011)

3. M. Koshiba, K. Saitoh, K. Takenaga, S. Matsuo, Multi-core fiber design and analysis: coupled-mode theory and coupled-power theory. *Opt. Express* **19**(26), B102–B111 (2011)
4. K. Takenaga et al., A large effective area multi-core fiber with an optimized cladding thickness. *Opt. Express* **19**(26), B543–B550 (2011)
5. K. Imamura, H. Inaba, K. Mukasa, R. Sugizaki, 19-core multi core fiber to realize high density space division multiplexing transmission, in *2012 IEEE Photonics Society Summer Topical Meeting Series* (2012), pp. 208–209
6. J. Sakaguchi et al., 19-core fiber transmission of $19 \times 100 \times 172$ -Gb/s SDM-WDM-PDM-QPSK signals at 305Tb/s, in *OFC/NFOEC* (2012), pp. 1–3
7. N. Kishi, E. Yamashita, A simple coupled-mode analysis method for multiple-core optical fiber and coupled dielectric waveguide structures. *IEEE Trans. Microw. Theory Tech.* **36**(12), 1861–1868 (1988)
8. S. Matsuo et al., 12-core fiber with one ring structure for extremely large capacity transmission. *Opt. Express* **20**(27), 28398–28408 (2012)
9. H.R. Stuart, Dispersive multiplexing in multimode optical fiber. *Science* **289**(5477), 281–283 (2000)
10. F. Saitoh, K. Saitoh, M. Koshiba, A design method of a fiber-based mode multi/demultiplexer for mode-division multiplexing. *Opt. Express* **18**(5), 4709–4716 (2010)
11. Y. Kokubun, M. Koshiba, Novel multi-core fibers for mode division multiplexing: proposal and design principle. *IEICE Electron. Express* **6**(8), 522–528 (2009)
12. X. Cen et al., “Supermodes in strongly-coupled multi-core fibers,” in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, 2013, pp. 1–3.
13. C. Xia et al., Hole-assisted few-mode multicore fiber for high-density space-division multiplexing. *IEEE Photonics Technol. Lett.* **24**(21), 1914–1917 (2012)
14. S. Matsuo, Y. Sasaki, I. Ishida, K. Takenaga, K. Saitoh, M. Koshiba, Multicore fiber with one-ring structure, in *Proceedings of the SPIE*, vol. 8647 (2013), p. 86470F
15. Z. Yibo et al., Supermode analysis in multi-core photonic crystal fiber laser, in *Proceedings of the SPIE*, vol. 7843 (2010), p. 784316
16. E. Kapon, J. Katz, A. Yariv, Supermode analysis of phase-locked arrays of semiconductor lasers. *Opt. Lett.* **9**(4), 125–127 (1984)
17. A. Szameit et al., Light evolution in arbitrary two-dimensional waveguide arrays. *Phys. Rev. A* **75**(5), 053814 (2007)
18. T. Mansuryan et al., Spatially dispersive scheme for transmission and synthesis of femtosecond pulses through a multicore fiber. *Opt. Express* **20**(22), 24769–24777 (2012)
19. J. Hudgings, L. Molter, M. Dutta, Design and modeling of passive optical switches and power dividers using non-planar coupled fiber arrays. *IEEE J. Quantum Electron.* **36**(12), 1438–1444 (2000)
20. K.L. Reichenbach, C. Xu, Numerical analysis of light propagation in image fibers or coherent fiber bundles. *Opt. Express* **15**(5), 2151–2165 (2007)
21. J. Zhou, Analytical formulation of super-modes inside multi-core fibers with circularly distributed cores. *Opt Express* **22**(1), 673–688 (2014). (in English). <https://doi.org/10.1364/Oe.22.000673>
22. Y.-C. Meng, Q.-Z. Guo, W.-H. Tan, Z.-M. Huang, Analytical solutions of coupled-mode equations for multiwaveguide systems, obtained by use of Chebyshev and generalized Chebyshev polynomials. *J. Opt. Soc. Am. A* **21**(8), 1518–1528 (2004)

23. C. Alexeyev, T.A. Fadeyeva, N.A. Boklag, M. Yavorsky, Supermodes of a double-ring fibre array with symmetric coupling. *Ukrainian J. Phys. Opt.* **12**, 83 (2011)
24. S. Peles, J.L. Rogers, K. Wiesenfeld, Robust synchronization in fiber laser arrays. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* **73**, 026212 (2006)
25. S.-Y. Tseng, Y. Kim, C.J.K. Richardson, J. Goldhar, Implementation of discrete unitary transformations by multimode waveguide holograms. *Appl. Opt.* **45**(20), 4864–4872 (2006)
26. A.W. Snyder, Coupled-mode theory for optical fibers. *J. Opt. Soc. Am.* **62**(11), 1267–1277 (1972)
27. W. Press, S. Teukolsky, W. Vetterling, B. Flannery, *Numerical Recipes In Fortran 77: the Art of Scientific Computing*, vol. 1 (1992), pp. 394–425
28. T. McMillen, On the eigenvalues of double band matrices. *Linear Algebra Appl.* **431**(10), 1890–1897 (2009)
29. A. Mafi, J.V. Moloney, Shaping modes in multicore photonic crystal fibers. *IEEE Photonics Technol. Lett.* **17**(2), 348–350 (2005)
30. J.H. Zhou, A non-orthogonal coupled mode theory for super-modes inside multi-core fibers. *Opt Express* **22**(9), 10815–10824 (2014). (in English). <https://doi.org/10.1364/Oe.22.010815>
31. J. Zhou, P. Gallion, A novel mode multiplexer/demultiplexer for multi-core fibers. *IEEE Photonics Technol. Lett.* **25**(13), 1214–1217 (2013)
32. B.E. Little, W.P. Huang, Coupled-mode theory for optical waveguides. *Prog. Electromagn. Res.* (1995)
33. W.-P. Huang, Coupled-mode theory for optical waveguides: an overview. *J. Opt. Soc. Am. A* **11**(3), 963–983 (1994)
34. M.G. Robert, Toeplitz and Circulant Matrices: A Review (2005)
35. J. Zhou, A non-orthogonal coupled mode theory for super-modes inside multi-core fibers. *Opt. Express* **22**(9), 10815–10824 (2014)
36. J.H. Zhou, Power splitting and switching in a multi-core fiber based on the multimode interference effect. *Opt Express* **23**(17), 22098–22107 (2015). (in English). <https://doi.org/10.1364/Oe.23.022098>
37. T. Watanabe, K. Suzuki, T. Takahashi, Silica-based PLC transponder aggregators for colorless, directionless, and contentionless ROADMs, in *OFC/NFOEC* (2012), pp. 1–3
38. J.M. Heaton, R.M. Jenkins, D.R. Wight, J.T. Parker, J.C.H. Birbeck, K.P. Hilton, Novel 1-to-N way integrated optical beam splitters using symmetric mode mixing in GaAs/AlGaAs multimode waveguides. *Appl. Phys. Lett.* **61**(15), 1754–1756 (1992)
39. W. Fan, Y. Jianyi, C. Limei, J. Xiaoqing, W. Minghua, Optical switch based on multimode interference coupler. *IEEE Photonics Technol. Lett.* **18**(2), 421–423 (2006)
40. J. Leuthold, J. Eckner, E. Gamper, P.A. Besse, H. Melchior, Multimode interference couplers for the conversion and combining of zero- and first-order modes. *J. Lightwave Technol.* **16**(7), 1228–1239 (1998)
41. J. Zhou, OAM states generation/detection based on the multimode interference effect in a ring core fiber. *Opt. Express* **23**(8), 10247–10258 (2015)

42. C. Brunet, B. Ung, P.A. Bélanger, Y. Messaddeq, S. LaRochelle, L.A. Rusch, Vector mode analysis of ring-core fibers: design tools for spatial division multiplexing. *J. Lightwave Technol.* **32**(23), 4648–4659 (2014)
43. Y. Xiaozhou et al., Primary experiments on 2-D and 1-D fiber-type optical phased array, in *Proceedings of the SPIE*, vol. 7136 (2008), p. 71363J

Chapter 5

Random Coupling Theory for Multi-mode Waveguides and Free Space Propagation



Abstract Random coupling theory for multi-mode waveguides and free space propagation. The chapter studies the random coupled mode theory, which has the coupling coefficients varies randomly and should be studied through the stochastic approach. The basic equations are given and the analytical results for the average value, as well as the 2nd and 4th order moments are studied analytically as well.

As is shown in the previous chapters, the coupling between the mode induces power transferring. What's more, the coupling effect might be random and either single mode waveguides or multimode waveguides suffer from such random refractive index variations induced coupling coefficient change. During the fabrication of the optical waveguides, the imperfect surface caused by etching resolution limit [1, 2] and the irregularity of the materials inside the waveguides will lead to random index variations either at the edges or within the cores of the waveguides. Such random index variations will introduce additional mode loss and mode coupling [1, 2]. As indicated by the literatures [1, 2], functional devices based on optical waveguides will have a degraded performance due to the effect. In this chapter, we introduce a random coupling theory for the waveguides, and it is later extended for the free space wave propagation. In the random coupling mode theory, the coupling coefficients varies randomly and should be studied through the stochastic approach. The basic equations are given and the analytical results for the average value, as well as the 2nd and 4th order moments are studied analytically as well.

5.1 A Comprehensive Analytical Model to Characterize Randomness in Optical Waveguides

Part of the section reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3].

In this section, the conventional coupled mode theory (CMT) is modified into the random CMT, which is based on the stochastic differential equations (SDEs)

by taking the random refractive index variations into account. The SDEs serves at the mathematical fundamentals, which can be used to derive the ordinary differential equations (ODEs) to analyze the modal amplitudes statistics. The higher order moments of modal amplitudes, such as the optical power, the power variations, and the power correlations can be computed by the ODEs. The ODEs have constant coefficients, and therefore, analytical solutions can be obtained. To validate the theory, the ODEs for mode power evolution are derived, which are in accordance with the Marcuse' coupled power equation. The higher order moments of the power, including the power variations and power correlations, are studied based on the proposed model, offering a more precise analysis compared with Marcuse' model. We conducted Monte-Carlo simulations to verify the proposed formulas.

5.1.1 Introduction

As is well known, the random index variations will affect the mode propagation inside the optical waveguides, and this is modelled by the coupled mode theory (CMT) [4]. Marcuse [5–7] derived a comprehensive theoretical framework based on the coupled mode equations, which has opened a door for the random index variations impact evaluation has resulted in consequent studies on this topic [1, 2, 8–12].

Despite the pioneering works [1, 2, 5–13] by Marcuse, there remain a handful of questions that have yet to be answered. Firstly, there is an absence of a comprehensive theoretical framework, which poses challenges in conducting a deeper analysis of the problem. Although the coupled power equation has been ingeniously derived by Marcuse, the precise equations for the higher order moments of the power, such as the power fluctuations and power correlations, which is essential in analysis, are not formulated. In [13], an approximate equation was provided by Marcuse, and it assumed that the mode powers were independent with respect to each other, which suggested the power correlation is 0. As is indicated by Marcuse [13], the approximation assumption was only valid under certain restrictions. It becomes inaccurate in a lossless two-mode waveguide, where the mode coupling effect indicates that the powers are significantly correlated. Hence, the power fluctuation computed by the model in [13] is not precise in the strong coupling regime. In addition to that, the power correlation has not been studied by Marcuse either [13]. These parameters play an essential role while evaluating the system performance, e.g., the power sensitivity and the capacity of the system. For instance, a batch of multimode waveguides, which are all made through an identical process, ought to exhibit comparable edge roughness. Nevertheless, owing to the randomness, the coupling of modes may result in different output mode power distributions for the different multimode waveguides, even when subjected to identical input conditions. A more pronounced case arises from the popular mode division multiplexing (MDM) systems, which adopt multimode waveguides for transmission. In such a scenario, random coupling effect will be crucial for channel characterization in optical multiple input multiple output (OMIMO) communication system [14–17]. In the deployed multimode fiber

links, the fiber index ought to be influenced by the temperature and strain fluctuations [18]. For instance, the temperature sensitivity of SiO₂ and Ge are 1.6×10^{-5} and $2 \times 10^{-3} \text{ }^\circ\text{C}$ [18], which are remarkable for a multimode fiber with the length of hundreds of kilometers. Similar to the free space optical communication systems [19], such random index variations along the propagation path give rise to the power fluctuations across the modes. The random power change is quantitated by the parameter of the scintillation index in free space communication system, which scales with the propagation distance [19]. Hence, it can be inferred that in the MDM systems, the varying output power scales with the fiber link distance and this will bring a more stringent receiver sensitivity requirement, particularly for the long-haul links. In addition to that, the channel covariance matrix, which is formed by the mode power correlations, is related to the channel capacity [17]. Therefore, the higher order moments of amplitude, including the power variance and the power correlation, are of great importance in the MDM systems.

To tackle the unresolved questions, the following works have been carried out: (1) we propose a set of stochastic differential equations (SDEs) as the mathematical fundamentals to study the mode propagation in randomly perturbed optical waveguides; (2) A general set of ordinary differential equations (ODEs) for the higher order moments of the mode amplitudes are derived based on the SDE; (3) Two specific sets of ODEs for the 2nd moments, i.e., the mode power, and the 4th order moments, i.e., the power variations and the power correlations, are formulated with analytical solutions. The derived ODEs for the powers agree exactly with the model derived by Marcuse; (4) Comprehensive Monte Carlo simulations are performed to compare with the theoretical prediction by the analytical formulas and a remarkable level of agreement has been attained.

5.1.2 Mathematical Model

5.1.2.1 The Stochastic Differential Equation Model

The mathematical formula continues from the first section, Chap. 3, which derived the coupled mode equation with index perturbation δn^2 . The index perturbation can be assumed to consist of two random independent processes, which are in the x - y plane and the z direction respectively, and one has the correlation for the index perturbation and the coupling coefficient $C_{mm'}(z)$ as:

$$\begin{aligned} \langle \delta n^2(x, y, z) \delta n^2(x', y', z') \rangle &= R_{xy}(x - x', y - y') \eta(z - z') \\ \langle C_{mm'}(z) C_{mm'}^*(z') \rangle &= \kappa_{mm'} \eta(z - z') \end{aligned} \quad (5.1)$$

where

$$\kappa_{mm'} = \frac{k_0^4}{16\omega^2\mu^2} \iint_{xy} \iint_{x'y'} \phi_m^*(x, y) \phi_m(x', y') \phi_{m'}(x, y) \phi_{m'}^*(x', y') R(x - x', y - y') dS dS' \quad (5.2)$$

In case the index perturbation occurs on a deterministic location on the x - y plane, e.g. the rough side wall of the waveguide, while it is still random along the z direction, the coupling coefficient $C_{mm'}(z)$ should be treated a one-dimensional random process $f(z)$ with a scaling factor.

$$\begin{aligned} C_{mm'}(z) &= \varsigma_{mm'} f(z) \\ \langle f(z) f(z') \rangle &= \eta(z - z') \end{aligned} \quad (5.3)$$

where $\varsigma_{mm'}$ is the scaling factor, which is related to the strength of the randomness, the rough surface location and the modal fields of mode m and m' , as is shown in Eq. (3.9). For instance, if the rough edge is at $x = a$, we have $\delta n^2 = \delta(x - a)f(z)$, and the scaling factor should be:

$$\varsigma_{mm'} = \frac{k_0^2}{j4\omega\mu} \int_{-\infty}^{\infty} \phi_m^*(a, y) \phi_{m'}(a, y) dy \quad (5.4)$$

With the random variables mentioned above, Eq. (3.6) is actually a set of Stratonovich sense SDEs, which can be numerically solved by taking the value in the middle of the integration step. In order to take the average of the terms for further analysis, the SDEs ought to be converted into the Ito sense SDE, which can be numerically solved by taking the value in the beginning of the integration step. We provide the detailed conversion process below.

As is mentioned earlier, Eq. (3.6) is in the Stratonovich sense, and hence, the RHS of the Equation is evaluated as:

$$da_m(z) = \sum_{m'} a_{m'} \left(z + \frac{dz}{2} \right) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \quad (5.5)$$

One may note that $C_{mm'}$ is a colored Gaussian noise [20–22], i.e. a Gaussian process filtered by a linear filter with the impulse response of

$$\mathbf{FT}^{-1} \left(\sqrt{\mathbf{FT}((C_{mm'}(z) C_{mm'}^*(z')))} \right) \quad (5.6)$$

where FT stands for the Fourier transform with $(z - z')$ as the variable and $\mathbf{FT}^{-1}$ stands for the inverse Fourier transform. This linear filtering can be achieved by a linear ordinary differential equation (ODE) system [20–22]. For instance, if

$$\langle C_{mm'}(z) C_{mm'}^*(z') \rangle = \frac{1}{2\alpha} \exp(-\alpha |z - z'|) \quad (5.7)$$

we have the ODE as:

$$\begin{aligned} \frac{dC_{mm'}}{dz} + \alpha C_{mm'} &= n(z) \\ \langle n(z)n(z') \rangle &= \delta(z - z') \end{aligned} \quad (5.8)$$

To convert the Stratonovich sense SDE to the Ito sense, we use the Taylor expansion [23]:

$$\begin{aligned} a_{m'}\left(z + \frac{dz}{2}\right) &= a_{m'}(z) + \frac{da_{m'}(z)}{dz} \frac{dz}{2} + \dots \\ &= a_{m'}(z) + \frac{da_{m'}(z)}{2} + \dots \end{aligned} \quad (5.9)$$

Combining Eq. (5.9) into Eq. (5.5), one has

$$da_m(z) = \sum_{m'} a_{m'}(z) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' + \sum_{m'} \frac{da_{m'}(z)}{2} \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \quad (5.10)$$

The Equation can be further converted into the reduced form by submitting da_m into Eq. (5.10), and one has

$$\begin{aligned} da_m(z) &= \sum_{m'} a_{m'}(z) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \\ &+ \sum_{m'} \frac{\sum_n a_n(z) \int_z^{z+dz} C_{m'n} e^{j(\beta_{m'} - \beta_n)z'} dz'}{2} \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \\ &+ O(dz^{3/2}) \end{aligned} \quad (5.11)$$

It is demonstrated in [13] that the second RHS term of Eq. (5.46) can be reduced into the compact form by letting $n = m$, because other terms will encounter fast oscillations in the z direction, which will be zero after averaging over z .

Therefore, one has

$$da_m(z) = \sum_{m'} a_{m'}(z) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz'$$

$$\begin{aligned}
& + \frac{1}{2} \left(\sum_{m'} \int_z^{z+dz} C_{m'm} e^{j(\beta_{m'} - \beta_m)z'} dz' \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \right) a_m(z) \\
& + O(dz^{3/2})
\end{aligned} \tag{5.12}$$

It is obvious that $C_{m'm} = C_{mm'}^*$, and one has:

$$\begin{aligned}
da_m(z) &= a_m(z + dz) - a_m(z) \\
&= \sum_{m'} a_{m'}(z) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' + \sum_{m'} \frac{da_{m'}(z)}{2} \int_z^{z+dz} C_{mm'} e^{j(\beta_{m'} - \beta_m)z'} dz' \\
&= \sum_{m'} a_{m'}(z) \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \\
&+ \frac{1}{2} \left(\sum_{m'} \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \int_z^{z+dz} C_{mm'}^* e^{-j(\beta_m - \beta_{m'})z'} dz' \right) a_m(z) \\
&+ O(dz^{3/2})
\end{aligned} \tag{5.13}$$

Equation (5.13) is the stochastic coupled mode equation, which is quite different from the one derived by Marcuse [5–7]. The second RHS term is the particular term which was not considered previously. As such, Eq. (5.13) is the mathematical fundamentals for the latter analysis.

5.1.2.2 Energy Conservation

To demonstrate the validity of Eq. (5.13), we will check the energy conservation law by the equation. The total power can be denoted as P_t

$$\langle P_t \rangle = \langle \mathbf{a}^H \mathbf{a} \rangle \tag{5.14}$$

where $\mathbf{a}$ is the vector with its i th element as a_i , and H stands for Hermitian transpose. Due to the fact that the coupling coefficient $C_{mm'}$ is the element of an anti-Hermitian matrix, one may conclude that

$$\frac{d\langle P \rangle_t}{dz} = 0 \tag{5.15}$$

which suggests that P_t is a constant. One should note that the powers of the leaky modes have been included in P_t . Otherwise, it will attenuate during the propagation.

5.1.2.3 Coupled Power Equation

The next step to demonstrate the validity of the model is to derive the coupled power equation. We start with the definition of the power as [5]

$$P_m(z) = \langle a_m(z) a_m^*(z) \rangle \quad (5.16)$$

Taking the differential of the term, one has:

$$d(a_m(z) a_m^*(z)) = d(a_m(z)) a_m^*(z) + a_m(z) d(a_m^*(z)) + da_m(z) da_m^*(z) \quad (5.17)$$

One may notice that the differential of SDE is different from that of ODE. The product of the differentials will not be zero after averaging, which is not the case in the differential of ODEs. In the Ito sense SDE, the random process is assumed to be Markovian, which indicates that a_m and δn^2 are independent variables. Hence, Combining Eqs. (5.13) and (5.17) and considering that the expectation of δn^2 and a_m are 0, we have:

$$\begin{aligned} dP_m(z) = & \sum P_{m'}(z) \left\langle \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \int_z^{z+dz} C_{mm'}^* e^{-j(\beta_m - \beta_{m'})z''} dz'' \right\rangle \\ & - \sum P_m(z) \left\langle \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \int_z^{z+dz} C_{mm'}^* e^{-j(\beta_m - \beta_{m'})z''} dz'' \right\rangle \end{aligned} \quad (5.18)$$

The RHS in Eq. (5.18) contains two terms, with one of them coming from the 2nd order differentials and the other term coming from the 1st order differentials in Eq. (5.17). One may further simplify the equation by using the formula below:

$$\left\langle \int_z^{z+dz} C_{mm'} e^{j(\beta_m - \beta_{m'})z'} dz' \int_z^{z+dz} C_{mm'}^* e^{-j(\beta_m - \beta_{m'})z''} dz'' \right\rangle = dz \kappa_{mm'} \int_{-dz}^{dz} \eta(z') e^{j(\beta_m - \beta_{m'})z'} dz' \quad (5.19)$$

where

$$\kappa_{mm'} = \frac{k_0^4}{16\omega^2 \mu^2} \iint_{xy} \iint_{x'y'} \phi_m^*(x, y) \phi_{m'}(x', y') R(x - x', y - y') dS dS' \quad (5.20)$$

In case that the step size dz is much larger than the random index correlation length, which is also assumed by Marcuse [5–7] and the subsequent researchers [1], Eq. (5.19) should be rewritten as,

$$\int_{-dz}^{dz} \eta(z') e^{j(\beta_m - \beta_{m'})z'} dz' \approx \mathbf{FT}(\eta)(\beta_m - \beta_{m'}) \quad (5.21)$$

where $\mathbf{FT}$ denotes Fourier transform. One will have:

$$\frac{dP_m}{dz} = \sum \kappa_{mm'} \mathbf{FT}(\eta)(\beta_m - \beta_{m'})(P_{m'}(z) - P_m(z)) \quad (5.22)$$

Equation (5.22) is in exact agreement with the coupled power equation derived by Marcuse under a totally different formulation [5–7]. In addition to that, Eq. (5.22) includes the randomness in the x - y plane, addressing a point missed in the equation derived by Marcuse. Hence, Eq. (5.22) is applicable not only for the waveguide sidewall roughness study, but also for the formulation of the beam propagation in the random media, e.g. free space beam propagation under the atmosphere turbulence.

5.1.2.4 Coupled Equation for the Higher Order Statistics

After the verification of the model by comparing with the existing theory, we consider more unexplored features. For instance, the mode power variations and their correlations are of great importance. Here, we demonstrate that the coupled OEDs can be derived for these terms based on the SDE.

We start with the definition for the optical power variance:

$$\langle P_m^2(z) \rangle = \langle a_m^2(z) a_m^{*2}(z) \rangle \quad (5.23)$$

Taking the differential on the term, one has

$$\begin{aligned} dP_m^2(z) &= 2a_m(z)a_m^{*2}(z)da_m(z) + 2a_m^*(z)a_m^2(z)da_m^*(z) \\ &\quad + 2a_m(z)2a_m^*(z)da_m(z)da_m^*(z) + a_m^{*2}(z)da_m(z)da_m(z) \\ &\quad + a_m^2(z)da_m^*(z)da_m^*(z) \end{aligned} \quad (5.24)$$

Combining Eqs. (5.13) and (5.24) and averaging the equation, one has:

$$\begin{aligned} \frac{d\langle P_m^2(z) \rangle}{dz} &= -2\langle P_m(z)^2 \rangle \sum_{m'} \kappa_{mm'} \mathbf{FT}(\eta)(\beta_m - \beta_{m'}) \\ &\quad + 4 \sum_{m'} \langle P_m(z) P_{m'}(z) \rangle \kappa_{mm'} \mathbf{FT}(\eta)(\beta_m - \beta_{m'}) \\ &\quad - 2\kappa_{mm} \mathbf{FT}(\eta)(0) \langle P_m(z)^2 \rangle \end{aligned} \quad (5.25)$$

There is a self-coupling term, i.e. the 3rd RHS term, in Eq. (5.22), which was not included in the approximated formula of coupled power fluctuation proposed by

Marcuse [13]. The term is not negligible as it is not absorbable due to the deterministic propagation constant, which does not cause oscillation. Nevertheless, the term can be combined with the other two terms and they cancel with each other. In addition to that, one has the power correlation term. Because one may not determine whether P_m and $P_{m'}$ are independent variables, one may not separate the power correlation term as the product of the powers, which was done in the approximated solution by Marcuse. In this case, one should derive the coupled equations for the power correlation terms by taking the differential of the term:

$$\begin{aligned}
 dP_m(z)P_n(z) = & da_m(z)a_m^*(z)a_n(z)a_n^*(z) + a_m(z)da_m^*(z)a_n(z)a_n^*(z) \\
 & + a_m(z)a_m^*(z)da_n(z)a_n^*(z) + a_m(z)a_m^*(z)a_n(z)da_n^*(z) \\
 & + da_m(z)da_m^*(z)a_n(z)a_n^*(z) + a_m(z)a_m^*(z)da_n(z)da_n^*(z) \\
 & + da_m(z)a_m^*(z)da_n(z)a_n^*(z) + a_m(z)da_m^*(z)a_n(z)da_n^*(z) \\
 & + da_m(z)a_m^*(z)a_n(z)da_n^*(z) + a_m(z)da_m^*(z)da_n(z)a_n^*(z) \quad (5.26)
 \end{aligned}$$

One may find the ODE by taking average on the equation:

$$\begin{aligned}
 \frac{d\langle P_m P_n \rangle}{dz} = & -\langle P_m P_n \rangle \sum_{m'} \kappa_{mm'} \mathbf{FT}(\eta) (\beta_m - \beta_{m'}) - \langle P_m P_n \rangle \sum_{n'} \kappa_{nn'} \mathbf{FT}(\eta) (\beta_n - \beta_{n'}) \\
 & + \sum_{m'} \langle P_m(z) P_{m'}(z) \rangle \kappa_{mm'} \mathbf{FT}(\eta) (\beta_m - \beta_{m'}) \\
 & + \sum_{n'} \langle P_n(z) P_{n'}(z) \rangle \kappa_{nn'} \mathbf{FT}(\eta) (\beta_n - \beta_{n'}) \\
 & - 2\kappa_{mn} \mathbf{FT}(\eta) (0) \langle P_m P_n \rangle \quad (5.27)
 \end{aligned}$$

One may compute the power standard deviations and the power correlation as

$$\begin{aligned}
 \sqrt{\delta P_m^2} &= \sqrt{\langle P_m^2 \rangle - \langle P_m \rangle^2} \\
 C(P_m, P_n) &= \langle P_m P_n \rangle - \langle P_m \rangle \langle P_n \rangle \quad (5.28)
 \end{aligned}$$

5.1.2.5 Analytical Solutions for the ODEs of Power and Power Variations

After derivation of the ODEs, the solutions will be explored. Since Eqs. (5.22), (5.25) and (5.27) are the ODEs with constant coefficients, analytical solutions exist. One may formulate Eq. (5.22) in the matrix form [5–7]:

$$\frac{d\mathbf{P}}{dz} = \mathbf{KP} \quad (5.29)$$

where $\mathbf{P}$ is a vector with its elements as the mode powers, and $\mathbf{K}$ is the coupling matrix whose elements are the coupling coefficients. The solution to Eq. (5.38) can be formulated as:

$$\mathbf{P}(z) = \exp(\mathbf{K}z)\mathbf{P}(0) \quad (5.30)$$

Similarly, one may rewrite Eqs. (5.25) and (5.27) in the matrix form:

$$\frac{d\mathbf{Q}(z)}{dz} = \mathbf{M}\mathbf{Q}(z) \quad (5.31)$$

where $\mathbf{Q}$ is a vector with its elements as the 2nd order moments of the powers,

$$\mathbf{Q} = \begin{pmatrix} \langle P_1 P_1 \rangle \\ \vdots \\ \langle P_1 P_n \rangle \\ \langle P_2 P_1 \rangle \\ \vdots \\ \langle P_2 P_n \rangle \\ \vdots \\ \langle P_n P_1 \rangle \\ \vdots \\ \langle P_n P_n \rangle \end{pmatrix} \quad (5.32)$$

One may solve the equation as:

$$\mathbf{Q}(z) = \exp(\mathbf{M}z)\mathbf{Q}(0) \quad (5.33)$$

One may note that there is no loss in the coupled equation, which suggests that the leaky modes have been considered. If the coupled equations only include the guided modes, the loss should be included due to the contribution from the leaky modes [5]. Such contributions can be observed in Eqs. (5.22), (5.25) and (5.27), if mode m' or mode n' are regarded as the radiation modes. Since the radiation modes attenuate remarkably with their power rapidly going to 0, their impact on other modes can be attributed solely to the radiation loss [5]. Henceforth, the radiation loss can be substituted by a deterministic loss term α_m as it is shown in Eqs. (5.22), (5.25) and (5.27), which can be calculated according to the mode field distribution and the parameters of the waveguide [5]. We will provide a thorough derivation for the loss term in the latter section.

5.1.2.6 Evaluation of the Modal Losses in Slab Waveguides and Planar Waveguides

In this part, the loss induced by the radiation modes will be evaluated analytically and the formulas will be compared with the published results.

For high index contrast waveguides, e.g. lithographically etched waveguides, the vector theory should be adopted to model the modes more accurately. Hence, the equation for the power coupling coefficients, i.e., Eq. (5.2), should be changed to the following:

$$\kappa_{mm'} = \frac{\omega^2 \varepsilon_0^2}{16} \iint_{xy} \iint_{x'y'} \mathbf{e}_m^*(x, y) \cdot \mathbf{e}_{m'}(x, y) \mathbf{e}_m(x', y') \cdot \mathbf{e}_{m'}^*(x', y') R(x - x', y - y') dS dS' \quad (5.34)$$

It is known that the plane wave function is the complete orthonormal basis. Therefore, the radiation modes $\mathbf{e}_{m'}$ can be expanded by the plane wave function and the integration should be over k_x and k_y to calculate the mode loss.

In the vector formulation, modes can be classified as the quasi TE and the quasi TM modes, whose z component of $\mathbf{E}/\mathbf{H}$ is small and negligible. Since the two types of modes are nearly orthogonal with each other, as D. Lenz did in [10], we use quasi TE modes in the analysis, which have the principal field components as E_x and H_y [10], while $E_x \gg E_y$. Hence, one may simplify Eq. (5.34) by this point, i.e., only the x components of $\mathbf{e}_m$ and $\mathbf{e}_{m'}$ will be included in the calculation [10]. The fabricated waveguides have four side walls, namely the side wall on the top, the bottom, the left and the right side. As an illustration, the mode loss due to the radiation by the top side wall will be derived with the rest three walls to be treated similarly. The actual loss will be the sum of the radiation losses from the four walls. We start the computation by assuming that the top wall is located at $y = b$, which suggests that the coupling coefficient in Eq. (5.34) will be

$$\kappa_{mm'} = \frac{\omega^2 \varepsilon_0^2}{16} \int_x \int_{x'} e_m^*(x, b) e_{m'}(x, b) e_m(x', b) e_{m'}^*(x', b) R_x(x - x') dx dx' \quad (5.35)$$

In the equation, e_m and $e_{m'}$ denote the x directional electrical fields. The plane waves are used to expand the radiation modes [10], and hence, $e_{m'}$ can be calculated as:

$$\begin{aligned} e_p(k_x, k_y) &= \rho_p(k_x, k_y) \exp(jk_x x + jk_y y) \\ \rho_p(k_x, k_y) &= \frac{1}{2\pi} \sqrt{\frac{2\omega\mu_0}{\beta_p(k_x, k_y)}} \\ \beta_p(k_x, k_y) &= \sqrt{n_2^2 k_0^2 - k_x^2 - k_y^2} \end{aligned} \quad (5.36)$$

where k_x and k_y represent the wave vectors in the x and y directions respectively, n_2 denotes the refractive index of the cladding, and β_p signifies the propagation constant of the plane wave. The plane wave function should be normalized such that the following is fulfilled [10].

$$\iint_{xy} e_p(k_x, k_y) e_p(k'_x, k'_y) dx dy = \frac{2\omega\mu_0}{\beta_p(k_x, k_y)} \delta(k_x - k'_x) \delta(k_y - k'_y) \quad (5.37)$$

Combining Eqs. (5.36) and (5.35), we may derive the loss coefficient as

$$\alpha_m = \int_{k_x} \int_{k_y} \kappa_{mp} \mathbf{FT}(\eta) (\beta_m - \beta_p) dk_x dk_y \quad (5.38)$$

$$\kappa_{mp} = \frac{\omega^2 \varepsilon_0^2 |\rho_p|^2}{16} \int_x \int_{x'} e_m^*(x, b) e_m(x', b) \exp(jk_x(x - x')) R_x(x - x') dx dx'$$

For the slab waveguide, e_m remains unaltered as x varies and therefore, the correlation function $R(x - x')$ will be a constant of $(n_1^2 - n_2^2)^2$, where n_1 represents the core index. Hence, one has $2\pi\delta(k_x)$ if the integration takes over $(x - x')$. Consequently, Eq. (5.39) is simplified as:

$$\alpha_m = \frac{\omega^2 \varepsilon_0^2 \omega \mu_0}{16\pi} (n_1^2 - n_2^2)^2 e_m^*(b) e_m(b) \int_{k_y} \mathbf{FT}(\eta) (\beta_m - \beta_p) \frac{1}{\sqrt{n_2^2 k_0^2 - k_y^2}} dk_y \quad (5.39)$$

In the equation, $e_m(b)$ represents the renormalized $e_m(x, b)$ in case the two dimensional mode $e_m(x, y)$ is converted into the one dimensional mode $e_m(y)$. We may assume the following variable changing rule,

$$k_y = n_2 k_0 \sin \theta \quad (5.40)$$

and the loss will be formulated as:

$$\alpha_m = \frac{\omega^2 \varepsilon_0^2 \omega \mu_0}{16\pi} (n_1^2 - n_2^2)^2 e_m^*(b) e_m(b) \int_0^\pi \mathbf{FT}(\eta) (\beta_m - n_2 k_0 \cos \theta) d\theta \quad (5.41)$$

which is in precise alignment with the outcomes in [9] provided that the normalization factor is properly selected.

For the 3D waveguides, e.g. the rectangular waveguides, it is usually assumed that the index is randomly perturbed in the z direction [10]. Therefore, the top wall, i.e. the plane $y = b$, will have random index variations identical in the x direction [10], and hence, one has $R(x - x') = (n_1^2 - n_2^2)^2$. With this formulation in Eq. (5.38), the mode loss is formulated as:

$$\alpha_m = \int_{k_x} \int_{k_y} \frac{\omega^2 \varepsilon_0^2 |\rho_p|^2 (n_1^2 - n_2^2)^2}{16} \left| \int_x e_m^*(x, b) \exp(jk_x x) dx \right|^2 \mathbf{FT}(\eta) (\beta_m - \beta_p) dk_x dk_y \quad (5.42)$$

which is in precise alignment with the outcomes in [10]. More generally, Eq. (5.38) should be implemented in case the index perturbation varies randomly simultaneously in the x and z directions, which is the case in practical waveguides produced by the lithography process. The induced correlation length is comparable with the z directional correlation length. Hence, one may overestimate the radiation loss by the model in [10] for the waveguides with short x/y directional correlation length.

For instance, Lenz et al. [10] discussed a rectangular waveguide with a core index of 1.01 and a cladding index of 1, which has the width and height both specified as 30 μm , as is mentioned in [10]. The autocorrelation function along the z-direction is a Gaussian function, featuring a correlation length of 0.75 μm , according to [10]. Furthermore, the standard deviation of the roughness $\bar{\sigma}$ is reported as 4×10^{-9} [10]. The signal wavelength considered in this context is 1550 nm. Based on Eq. (5.42), the radiation loss is 0.0006 Np/m, or 0.005 dB/m for the E_x^{11} mode, which is in alignment with the outcomes [10]. However, the radiation loss is different if the two-dimensional random process model is implemented. Figure 5.1 illustrates the variation in radiation loss as the correlation length in the x/y direction changes, while all other parameters remain constant.

As is evident from Fig. 5.1, the radiation loss escalates with an increase in the correlation length in the x/y direction, while keeping other parameters constant. When comparing scenarios where the x/y directional correlation length is infinitely long versus when it equals the z-directional correlation length (i.e., 0.75 μm), a noticeable discrepancy of 0.0007 dB/m (15%) is observed. However, in cases where the z-directional correlation length is significantly shorter, for instance, 0.05 μm as reported in [11], the discrepancy between these two scenarios can soar up to 80%, as depicted in Fig. 5.2 (comparing results with a correlation length of 0.05 μm to those with an infinite correlation length). Consequently, the findings presented in [10] can be considered reliable primarily for waveguides exhibiting large correlation lengths of roughness. For waveguides fabricated using advanced lithography processes, which often yield roughness with correlation lengths as short as 50 nm, Eq. (5.38) should be employed to procure a more accurate estimation of the radiation loss.

5.1.3 Simulation Results and Discussion

We have conducted numerical simulations to validate the proposed analytical formulas. A slab waveguide identical to the one in Marcuse' section [5] is selected as an example, which possesses the core index of 1.5 and the core to cladding index ratio of 1.01. The waveguide width is assumed to be 2d with $k_0 d = 82$ as indicated

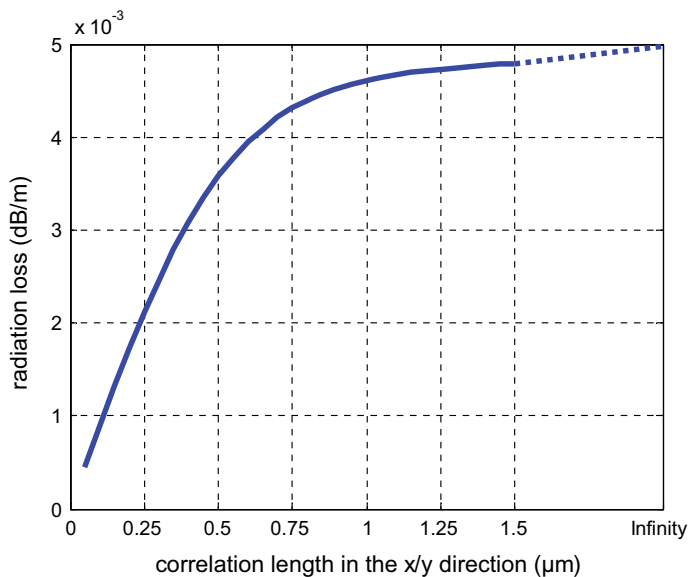


Fig. 5.1 The radiation loss of E_x^{11} mode versus correlation length in the x/y direction, the z directional correlation length is $0.75 \mu\text{m}$. This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

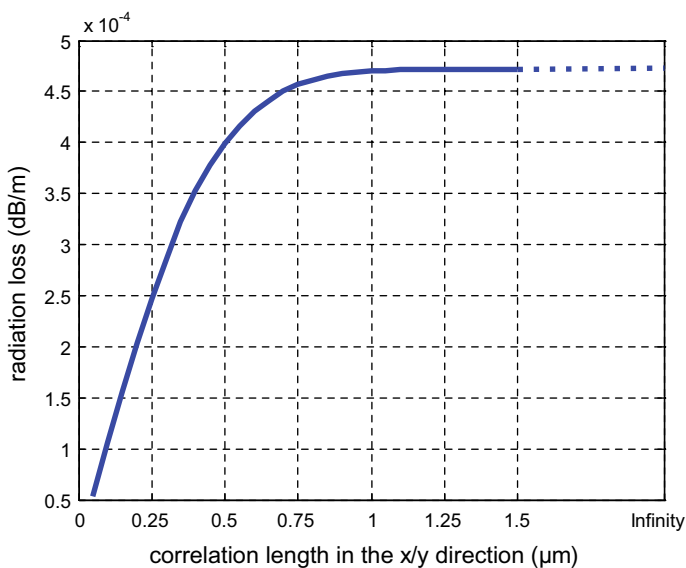


Fig. 5.2 The radiation loss of E_x^{11} mode versus correlation length in the x/y direction, the z directional correlation length is 50 nm . This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

Table 5.1 The effective indexes of the ten modes of the waveguide

Mode order	Effective index
1	1.499891
2	1.499563
3	1.499017
4	1.498254
5	1.497275
6	1.496082
7	1.494678
8	1.493068
9	1.49126
10	1.489269

This table is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

in [5] to specify the relation between the wavelength and waveguide parameter. The signal wavelength is assumed to be 1550 nm. The optical waveguide is computed to be able to transmit ten modes [5], with the effective indexes of the ten mode listed in Table 5.1.

Two random walls located at $x = d$ and $x = -d$ are assumed, which have the two independent random processes $f(z)$ and $h(z)$ to model the random index variations [5] caused by the fabrication process, or by the temperature and the strain variations.

One may compute the coupling coefficient $C_{mm'}$ by Eqs. (3.9) and (5.4) as follows [5]:

$$\begin{aligned}
 C_{mm'}(z) &= \frac{(n_1^2 - n_2^2)k_0^2 \phi_m \phi_{m'} \sqrt{\gamma_m \gamma_{m'}}}{2j\sqrt{\beta_m \beta_{m'}}(1 + \gamma_m d)(1 + \gamma_{m'} d)} \left(f(z) - (-1)^{m+m'} h(z) \right) \\
 \phi_m &= \begin{cases} \cos(\kappa_m d) & m = 0, 2, 4 \dots \\ \sin(\kappa_m d) & m = 1, 3, 5 \dots \end{cases} \\
 \gamma_m &= \sqrt{\beta_m^2 - n_2^2 k_0^2} \\
 \kappa_m &= \sqrt{n_1^2 k_0^2 - \beta_m^2}
 \end{aligned} \tag{5.43}$$

where n_1 and n_2 represent the core and cladding indexes, β_m stands for the m th mode's propagation constant, and ϕ_m signifies the m th mode field at $x = d$.

The correlation function of the two random processes are assumed to be the Gaussian function with the same correlation length [5]

$$\langle f(z)f(z') \rangle = \langle h(z)h(z') \rangle = \eta(z - z') = \bar{\sigma}^2 e^{-\frac{(z-z')^2}{D^2}} \tag{5.44}$$

where $\bar{\sigma}^2$ stands for the roughness strength, and D is the correlation length of the two processes.

The correlation function has its Fourier transform as [5]

$$\mathbf{FT}\left(\bar{\sigma}^2 e^{-\frac{(z-z')^2}{D^2}}\right) = \bar{\sigma}^2 \sqrt{\pi} D e^{-\frac{D^2 \omega^2}{4}} \quad (5.45)$$

One may calculate the correlation function of the coupling coefficient as [5]

$$\langle C_{mm'}(z) C_{mm'}^*(z') \rangle = \frac{(n_1^2 - n_2^2)^2 k_0^4 \phi_m^2 \phi_{m'}^2 \gamma_m \gamma_{m'}}{2 \beta_m \beta_{m'} (1 + \gamma_m d) (1 + \gamma_{m'} d)} \bar{\sigma}^2 e^{-\frac{(z-z')^2}{D^2}} \quad (5.46)$$

The radiation loss of each mode can be evaluated accordingly. Without loss of generality, we directly use Marcuse's formula [5]

$$\begin{aligned} \alpha_m &= \int_{-n_2 k_0}^{n_2 k_0} \mathbf{FT}(\eta) (\beta_m - \beta) I_m(\beta) d\beta \\ I_m(\beta) &= \frac{(n_1^2 - n_2^2) k_0^3 n_1 \sin^2 \theta_m}{2 \pi d \cos \theta_m (1 + \gamma_m d)} \\ &\quad \left(\frac{\rho \cos^2 \sigma d}{\rho^2 \cos^2 \sigma d + \sigma^2 \sin^2 \sigma d} + \frac{\rho \sin^2 \sigma d}{\rho^2 \sin^2 \sigma d + \sigma^2 \cos^2 \sigma d} \right) \\ \rho &= \sqrt{n_2^2 k_0^2 - \beta^2} \\ \sigma &= \sqrt{n_1^2 k_0^2 - \beta^2} \\ \sin \theta_m &= \frac{(m+1)\pi}{2 n_1 k_0 d} \end{aligned} \quad (5.47)$$

In case the correlation length $D \gg d$ or $D \ll d$, one has the approximation for the radiation loss. For example, When $D \ll d$, one has [5]

$$\alpha_m = \frac{\sqrt{\pi} (n_1^2 - n_2^2) k_0^3 n_1 \sin^2 \theta_m}{2 \cos \theta_m \left(1 + \frac{1}{\gamma_m d}\right)} \bar{\sigma}^2 \frac{D}{d} \quad (5.48)$$

When $D \gg d$, one has

$$\alpha_m = \bar{\sigma}^2 e^{-\frac{D^2}{4} (\beta_m - n_2 k_0)^2} \frac{n_1 k_0^3 \sin^2 \theta_m}{2 d \sqrt{n_2 k_0 (\beta_m - n_2 k_0)} \left(1 + \frac{1}{\gamma_m d}\right) \cos \theta_m}$$

$$\left\{ \begin{array}{ll} \frac{2n_2}{k_0 D^2 (\beta_m - n_2 k_0)} (\cot^2 \sigma d + \tan^2 \sigma d) & \text{and } \begin{array}{l} \tan \sigma d \neq 0 \\ \cot \sigma d \neq 0 \end{array} \\ (n_1^2 - n_2^2) & \text{or } \begin{array}{l} \tan \sigma d = 0 \\ \cot \sigma d = 0 \end{array} \end{array} \right. \quad (5.49)$$

We implement the analytical model derived in Sect. 5.2 to calculate the power, the power variation and the power correlations with the loss term included. Monte-Carlo simulation is conducted based on the SDE to check the accuracy of the analytical formulas.

First of all, it is assumed that D is $35d$ [5] and only the fundamental mode is injected with the unit power. We implement the normalized propagation distance and normalized power, with the distance normalized by the following [5]

$$z \frac{\bar{\sigma}^2 k_0^2}{d^2} \quad (5.50)$$

The normalization is applied to the power at each mode by the fundamental mode power. Ensuring the accuracy during numerical integration, it is necessary to adopt small propagation steps. Consequently, performing Monte Carlo simulations becomes a time-intensive process. In this simulation, the normalized propagation distance is constrained to 16,000 units, and a comprehensive set of 4000 simulation rounds are executed on the stochastic differential equation (SDE) to attain reliable average values.

Figures 5.3 and 5.4 depict the evolution of power in modes 1 through 4 as a function of the normalized propagation distance. Due to page constraints, not all modes are presented in the figures, but the output powers of all modes at the end of the propagation are summarized in Table 5.2. The Monte Carlo simulation relies on integrating the SDE given in Eq. (5.13), while the analytical results are based on the ODEs, e.g. Eq. (5.29). The curves from the Monte Carlo simulation exhibit good agreement with the analytical predictions. Evidently from the figures, the coupling between the fundamental mode and mode 2 is the most pronounced. As the mode order increases, the coupling strength diminishes, which can be attributed to the fact that coupling coefficients are influenced by the differences in propagation constants. The coefficients become smaller with greater propagation constant differences. Furthermore, it is observable that as the propagation distance increases, the average power in each mode tends to equalize, which is consistent with Marcuse's theory [5]. The accuracy of the analytical coupled power model is assessed against the Monte Carlo simulation. The relative error between the results remains below 5%, affirming the applicability of the proposed SDEs (i.e., Eq. (5.13)) in this study.

Subsequently, the higher-order statistics of the modal amplitudes were computed. Figures 5.5 and 5.6 illustrate the power standard deviations of modes 1 through 4 as a function of the normalized propagation distance. Table 5.2 summarizes the power standard deviations of each mode at the conclusion of the propagation. It can be confidently stated that the analytical solution of the ODEs presented in Eq. (5.33)

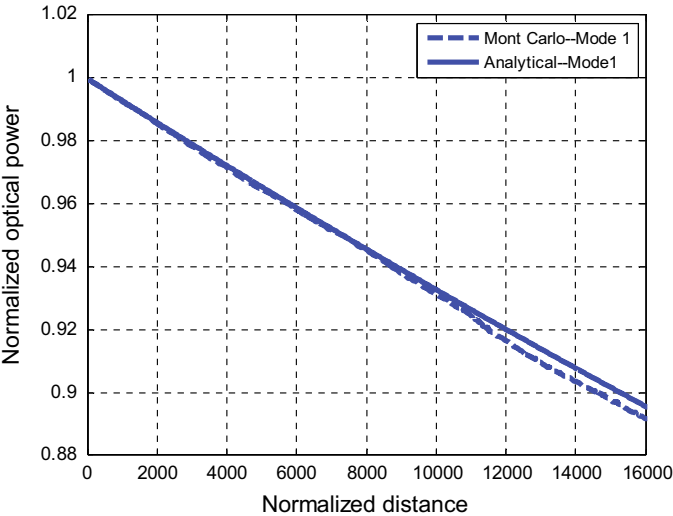


Fig. 5.3 The power evolution of the fundamental mode (mode 1). This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

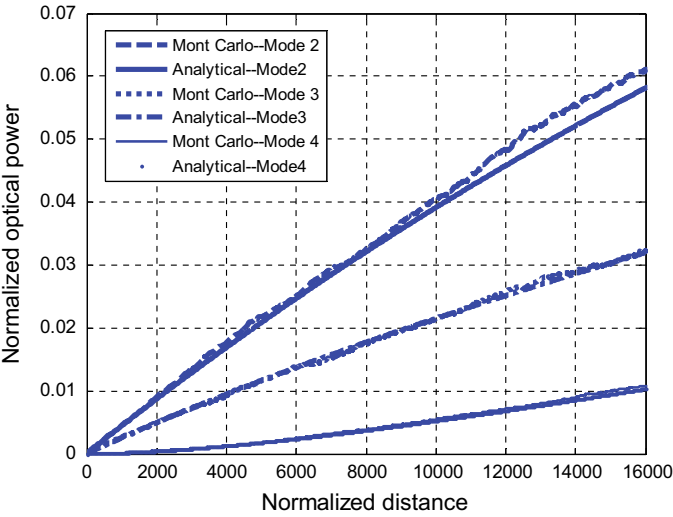


Fig. 5.4 The power evolution of mode 2 to mode 4. This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

Table 5.2 The output power and output power standard deviations at the end of the propagation

Mode order	Power after propagation of 16,000 normalized units		Power standard deviation after propagation	
	Analytical	Monte-Carlo	Analytical	Monte-Carlo
1	0.8943	0.8907	0.0689	0.0720
2	0.0588	0.0613	0.0565	0.0593
3	0.0323	0.0327	0.0326	0.0339
4	0.0106	0.0110	0.0128	0.0126
5	0.0032	0.0034	0.0049	0.0052
6	0.0007	0.0008	0.0013	0.0013
7	0.0001	0.0001	0.0002	0.0002
8	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000
10	0.0000	0.0000	0.0000	0.0000

This table is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

provides highly accurate predictions for these higher-order statistics, with a relative error remaining below 5%. This underscores the reliability and precision of the proposed analytical approach.

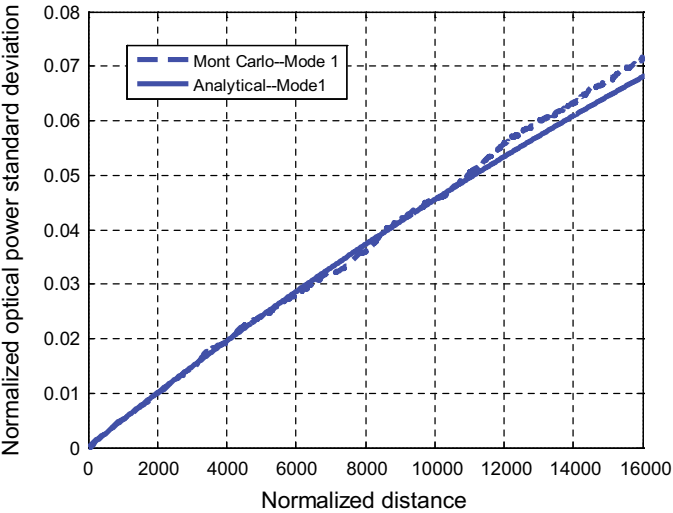


Fig. 5.5 The power standard deviation evolution of the fundamental mode. This figure is reprinted/ adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

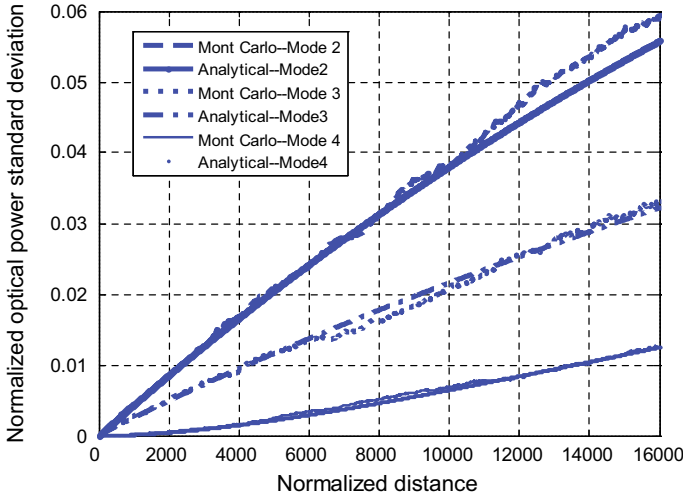


Fig. 5.6 The power standard deviation evolution of mode 2 to mode 4. This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

Given that the optical signal is deterministic at the start of propagation, the standard deviation of the mode power inherently starts at 0. Consequently, it is not possible to evaluate the normalized correlation coefficient as per Eq. (5.28), where normalization involves dividing by these standard deviations. Instead, we calculate the un-normalized power correlation coefficients between the powers of the fundamental mode and those of modes 2 through 4 by Eq. (5.28). These coefficients are then plotted in Fig. 5.7, revealing a trend: as the propagation distance lengthens, the correlation between the powers of these modes intensifies.

The novel insight into the higher-order statistics of modal amplitudes, which has not been previously explored, is invaluable for assessing system performance, as highlighted in preceding sections. As is evident from Figs. 5.5 and 5.6, the power variation escalates with increasing propagation distance, posing a challenge in receiver design due to sensitivity penalties. Furthermore, Fig. 5.7 demonstrates that the correlation, quantified by the absolute value of correlation coefficients, intensifies with propagation distance. Such a variation in correlation coefficients arises from random index fluctuations caused by temperature or strain within a single waveguide could lead to a reduction in system capacity, underscoring the importance of understanding and mitigating these effects.

To validate the robustness of the analytical findings, the correlation length D was varied within a wide range, spanning from $0.01d$ to $100d$ [5]. Despite this extensive variation, the deviation between the analytical solution of the ODEs and the Monte Carlo simulation remained consistently below 5%. This underscores the robustness of the model presented in this section, effectively demonstrating its capability to

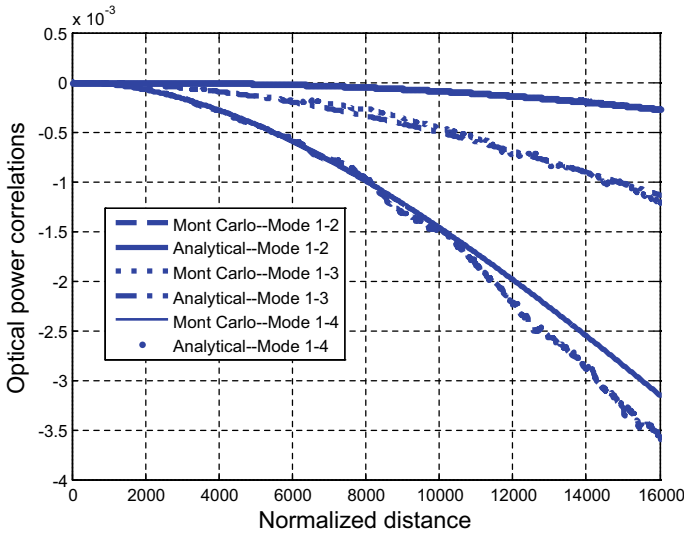


Fig. 5.7 The power correlation coefficients between mode 1 and mode 2 to mode 4 along the propagation distance. This figure is reprinted/adapted with permission from [Opt. Express 24, 6825–6842 (2016)] © The Optical Society of America [3]

simulate random optical waveguides characterized by diverse correlation lengths with high accuracy.

It is noteworthy that the analytical model offers a significant enhancement in computational efficiency compared to Monte-Carlo simulations. Specifically, the computational time required is reduced by an impressive factor of 10,000, highlighting the substantial advantages of the proposed analytical approach in terms of speed and practicality.

5.1.4 Summary

Utilizing the coupled mode theory, a series of SDEs have been formulated to investigate random index fluctuations within optical waveguides. Correspondingly, ODEs have been derived to describe both the power evolution and higher-order statistics of modal amplitudes. These ODEs, featuring constant coefficients, allow for analytical solutions, thereby significantly reducing computational demands and eliminating the need for time-consuming Monte Carlo simulations. Moreover, they offer a profound insight into the effects of random index variations.

To validate the derived ODEs, extensive Monte Carlo simulations have been conducted, focusing on a slab waveguide as a case study proposed by Marcuse [5]. However, it is anticipated that these ODEs will remain applicable and functional for optical waveguides of arbitrary geometries. This approach not only paves way for the

analysis but also enhances our understanding of the intricate relationships between random index variations and waveguide performance.

5.2 Analytical Study on the Evolutionary Behavior of Transfer Matrix Element Moments in Strongly Coupled Multimode Systems

Part of the section reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24].

In this section, the random coupled mode theory has been used to evaluate the evolutionary behavior of the transfer matrix element moments in strongly coupled multimode systems. Using the similar derivation process in Sect. 5.1, the randomly coupled multimode system is modeled by a set of coupled stochastic differential equations (SDEs), which can be used to find the coupled ordinary differential equations (ODEs) for the arbitrary order moments of the matrix elements. Analytical solutions to the ODEs are derived, which made it possible to analyze the asymptotic behaviors of the solutions. Comparing with the existing results derived from the property of the Haar matrix, the proposed theory exhibits a perfect match. Monte Carlo simulations are performed to check the evolutionary behaviors of the transfer matrix element moments, obtained through the analytical approach.

5.2.1 Introduction

Multimode waveguides, notably multimode fibers, have gained significant attention lately owing to their promising potential across diverse fields, including communications and sensing technologies [25–29]. Within multimode waveguides, the modes undergo coupling with each other as a result of random index fluctuations. The strength of this coupling can vary from weak to strong, determined by the magnitude of the fluctuations and the fulfillment of the phase matching condition.

As shown in the previous section, the random coupling effect results in a transfer matrix [30], which has random matrix elements. The moments of the elements can be used for the analysis of the average power, the power variance and the impact of the fiber nonlinear effect analysis [30–34]. While the former two applications have been elaborated in the previous section, we discuss the third application in the following. While numerically integrating the nonlinear Schrödinger equations (NLSE), it is necessary to average the nonlinear term [30–34]. Particularly, the fourth order moments of the mode amplitudes [30] remarkably impact the evaluation of the nonlinear term. Utilizing the statistical properties of the fourth-order matrix element moments, the crucial propagation equation governing nonlinear coupling within multimode fibers, known as the generalized Manakov equation, has been

derived [30–34]. The multimode Raman amplification process requires the employment of fourth-order transfer matrix element moments for precise analysis as well, as evidenced in Refs. [35, 36]. Specifically, the nonlinear coupling coefficients Q_{plmn} in Eq. (6) of Ref. [35] play a crucial role in modal amplification characterization. The averaged value of these coefficients is directly linked to the fourth-order moments of the transfer matrix elements. The third example demonstrating the application of statistical information derived from transfer matrix moments for nonlinearity analysis is in the multimode Brillouin sensing systems [37], where random coupling plays a pivotal role in redistributing the modal-dependent Brillouin gain. To illustrate, in a fiber featuring two degenerate linearly polarized modes, the Brillouin gain exhibits a strong dependence on the transfer matrix element moments, as demonstrated in [38]. The examination of transfer matrix element moments is also crucial for characterizing modal dispersion [39–41], a vital parameter in mode division multiplexing systems and multimode sensing systems [42–46].

In multimode fibers, the coupling dynamics exhibit such a rule: weak coupling between different modal groups and strong coupling within the same modal group [42–46]. This strong intra-group coupling induces a random nature to the transfer matrix, attracting special attention during analysis [31–34, 39, 40, 42, 45]. Currently, many studies on strongly coupled multimode systems assume a purely random transfer matrix which is uniformly distributed [30–34]. Under this assumption, the transfer matrix can be considered a Haar matrix, enabling the analytical derivation of its moments [30, 47]. This approach holds valid after the modes have transmitted for a sufficiently long distance within the strongly coupled mode system.

Despite the remarkable progress made in the previous works, a crucial question remains unanswered: how do the matrix element moments evolve in strongly coupled mode systems before reaching their asymptotic values? For relatively short multimode fibers, this evolutionary statistical information becomes indispensable, particularly in applications such as multimode parametric amplifiers with short, highly nonlinear fibers and sensing systems with short fiber lengths. Even for systems with longer fibers, understanding the modal power evolutionary behavior at the fiber's inception is intriguing. While Monte Carlo simulations can indeed reveal the evolutionary behaviors of transfer matrix element moments [30], they are often time-consuming and lack direct physical insights.

In our recent research [48], we introduced a stochastic differential equation (SDE) approach to investigate Mueller matrix element moments in polarization state analysis. Here, we extend this methodology to study the evolutionary behaviors of transfer matrix element moments in multimode waveguides/fibers. By establishing a fundamental SDE for the transfer matrix, we systematically derive SDEs for arbitrary-order moments of its elements. Subsequently, ordinary differential equations (ODEs) are derived by averaging both sides of the SDEs, effectively avoiding the need for numerous Monte Carlo simulations. If the random coupling strength remains constant along the propagation direction, the ODE coefficients remain fixed, enabling the derivation of the analytical solutions. This framework allows for an analytical exploration of the evolutionary behavior of random mode coupling effects.

The subsequent sections are structured as follows: Sect. 5.2 outlines the fundamental theory of SDEs and presents our key findings, particularly the derived ODEs governing the second-order and fourth-order moments of the transfer matrix elements. Section 5.3 discusses the asymptotic behaviors of these moments, demonstrating their alignment with established literature [30, 47]. Section 5.4 presents the Monte Carlo simulation results, demonstrating a close alignment with the analytical solutions. Section 5.5 concludes this section by summarizing its key takeaways.

5.2.2 Theory

Instead of the mode amplitude vector, we use the transfer matrix to derive the corresponding equations [39] for the random mode coupling modes. Assuming a total of N linearly polarized (LP) modes exist within the multimode fiber, where each of these modes is further divided into two polarized sub-modes, resulting in a total of $2N$ modes. The propagation of these modes within randomly coupled multimode fibers can be comprehensively described by [39]

$$|\psi(z)\rangle = \mathbf{U}|\psi(0)\rangle, \quad (5.51)$$

where $|\psi(z)\rangle$ is the mode amplitude vector, $\mathbf{U}$ is the transfer matrix relating the input and output vectors. It is derived that [39, 41]:

$$\frac{d\mathbf{U}}{dz} = j\frac{1}{2N}\left(\beta_0\mathbf{I} + \left(\vec{\beta} \cdot \vec{\Lambda}\right)\right)\mathbf{U}, \quad (5.52)$$

where $\beta_0\mathbf{I}$ represents for the general phase term, vector $\vec{\beta}$ is generalized birefringence vector with $4N^2 - 1$ elements, which includes the birefringence vectors of the N propagation modes and the coupling coefficients of the modes, $\vec{\Lambda}$ is the matrix vector comprises $4N^2 - 1$ elements, where each element is a $2N \times 2N$ matrix. These matrices characterize two distinct types of couplings within the multimode fiber, i.e., the coupling between polarized modes within one LP mode, and the coupling between the different LP modes. The Gell-Mann matrix [40] may be chosen as the matrix basis to form $\vec{\Lambda}$. However, we have adopted a slightly different formulation in Eq. (5.52).

The $4N^2 - 1$ elements of $\vec{\beta}$ are modeled as non-correlated Gaussian random processes, all sharing an identical distribution, as outlined in [39]. This uncorrelated white noise assumption for the generalized birefringence vector $\vec{\beta}$ is a prevalent approach in the research community studying strongly coupled multimode systems, as evidenced by numerous works such as [39–41, 49, 50]. Analogous treatments have been applied in analyzing the impact of polarization mode dispersion (PMD), which has been experimentally validated [51]. While the term $\beta_0\mathbf{I}$ represents a common phase term, it is typically omitted in the evaluation of matrix element moments due

to its insignificance [39]. In this work, it is retained and combined with $\vec{\beta}$. Similarly, the identity matrix $\mathbf{I}$ is combined with the matrix vector $\vec{\Lambda}$, so that the formulation can be simplified

$$\begin{aligned}\vec{\beta}' &= \begin{pmatrix} \beta_0 \\ \vec{\beta} \end{pmatrix} \\ \vec{\Lambda}' &= \begin{pmatrix} \mathbf{I} \\ \vec{\Lambda} \end{pmatrix},\end{aligned}\tag{5.53}$$

Hence, we may convert Eq. (5.52) into

$$\begin{aligned}\frac{d\mathbf{U}}{dz} &= j\mathbf{T}\mathbf{U} \\ \mathbf{T} &= j\frac{1}{2N}\left(\beta_0\mathbf{I} + \left(\vec{\beta} \cdot \vec{\Lambda}\right)\right) = \frac{1}{2N}\vec{\beta}' \cdot \vec{\Lambda}',\end{aligned}\tag{5.54}$$

without loss of generality, we may assume β_0 as the non-correlated Gaussian random process with respect to the elements of $\vec{\beta}$, sharing the same random distribution. Therefore, one has

$$\begin{aligned}\langle \beta'_m(z)\beta'_n(z') \rangle &= \frac{\sigma^2}{4N^2}\delta_{mn}\delta(z-z') \\ \langle \vec{\beta}' \cdot \vec{\beta}' \rangle &= \sigma^2\delta(z-z'),\end{aligned}\tag{5.55}$$

where σ^2 is the variance, which quantitates the mean-square magnitude of the generalized birefringence parameter γ [39]. In [39], it is formulated that $\langle \vec{\beta} \cdot \vec{\beta} \rangle = \omega^2\gamma^2\delta(z-z')$, where ω represents the angular frequency deviation from the central wavelength of the reference signal, while γ denotes the mean-square magnitude of the generalized birefringence parameter. This γ parameter is analogous to the polarization mode dispersion (PMD) value observed in birefringent fibers. It is important to clarify that when ω equals zero, it does not imply a complete absence of randomness or fluctuations. Rather, it signifies that the signal propagates within the same spatial rotational reference frame as the signal at the center frequency. This indicates that, at this specific frequency point, the signal experiences minimal or no change in its polarization state relative to the reference frame defined by the center frequency signal. Alternatively, one may opt to set the reference signal frequency to zero, as done in [23, 52] for the purpose of analyzing the randomness induced by polarization mode dispersion (PMD). In this context, ω represents the optical angular frequency. With the above-mentioned formula and Eq. (5.55), we have,

$$\sigma^2 = \frac{\omega^2\gamma^2 4N^2}{4N^2 - 1},\tag{5.56}$$

According to Eqs. (5.54 and 5.55), it is shown that $\mathbf{T}$ has its elements to be non-correlated identically distributed Gaussian processes,

$$\langle t_{mn}(z)t_{m'n'}^*(z') \rangle = \delta_{mm'}\delta_{nn'} \frac{\sigma^2}{8N^3} \delta(z - z'), \quad (5.57a)$$

$$\langle \mathbf{T} d\mathbf{z} \mathbf{T} d\mathbf{z} \rangle = \langle \mathbf{T} d\mathbf{z} \mathbf{T}^H d\mathbf{z} \rangle = \frac{\sigma^2}{4N^2} \mathbf{I} d\mathbf{z}, \quad (5.57b)$$

We have provided the detailed proof of the equation in the Appendix.

As is stated in the previous section, Eq. (5.54) is the Stratonovich sense SDE, suggesting that the integration takes place in the middle of the step $[z, z + dz]$, i.e. the point of $z + dz/2$,

$$d\mathbf{U} = j\mathbf{T} d\mathbf{z} \mathbf{U} \left(z + \frac{dz}{2} \right), \quad (5.58)$$

The Stratonovich sense SDE is generally converted to the Ito sense SDE, which has the integration taken place in the beginning of the step:

$$\begin{aligned} d\mathbf{U} &= j\mathbf{T} d\mathbf{z} \mathbf{U}(z) - \frac{1}{2} \langle \mathbf{T} d\mathbf{z} \mathbf{T} d\mathbf{z} \rangle \mathbf{U}(z) \\ &= j\mathbf{T} d\mathbf{z} \mathbf{U} - \frac{\sigma^2}{8N^2} \mathbf{I} d\mathbf{z} \mathbf{U}, \end{aligned} \quad (5.59)$$

The derivation of the equation is like the procedure in [49, 50], which is shown below.

The Stratonovich sense SDE should be expanded by the Taylor series

$$\begin{aligned} d\mathbf{U} &= j\mathbf{T} \mathbf{U} \left(z + \frac{dz}{2} \right) dz \\ &= j\mathbf{T} d\mathbf{z} \mathbf{U}(z) + j\mathbf{T} d\mathbf{z} \frac{d\mathbf{U}}{2} + O(dz), \end{aligned} \quad (5.60)$$

Since

$$\begin{aligned} j\mathbf{T} d\mathbf{z} \frac{d\mathbf{U}}{2} &= j\frac{1}{2} \mathbf{T} d\mathbf{z} \left(j\mathbf{T} d\mathbf{z} \mathbf{U}(z) + j\mathbf{T} d\mathbf{z} \frac{d\mathbf{U}}{2} + O(dz) \right) \\ &= -\frac{1}{2} \mathbf{T} d\mathbf{z} \mathbf{T} d\mathbf{z} \mathbf{U}(z) + O(dz), \end{aligned} \quad (5.61)$$

One has

$$d\mathbf{U} = j\mathbf{T} d\mathbf{z} \mathbf{U}(z) - \frac{1}{2} \mathbf{T} d\mathbf{z} \mathbf{T} d\mathbf{z} \mathbf{U}(z) + O(dz). \quad (5.62)$$

By neglecting the higher-order term denoted as $O(dz)$, the aforementioned equation can be transformed to align with the Ito interpretation in the mean-square limit sense, i.e.

$$d\mathbf{U} = j\mathbf{T}dz\mathbf{U}(z) - \frac{1}{2}\langle\mathbf{T}dz\mathbf{T}dz\rangle\mathbf{U}(z), \quad (5.63)$$

which is Eq. (5.59).

Equation (5.59) serves as the cornerstone for the subsequent developments in this work, providing the stochastic differential equation (SDE) that governs the system dynamics. Building upon this foundation, we can derive the SDE for the second-order moments, which capture essential statistical properties of the system. Specifically, in the mean-square limit sense, we obtain Eq. (5.64) as a starting point for analyzing these higher-order statistics.

$$d\mathbf{U} \otimes d\mathbf{U}^* = \langle\mathbf{T}dz \otimes \mathbf{T}^*dz\rangle(\mathbf{U} \otimes \mathbf{U}^*), \quad (5.64)$$

Therefore, we have:

$$\begin{aligned} d(\mathbf{U} \otimes \mathbf{U}^*) &= d\mathbf{U} \otimes \mathbf{U}^* + \mathbf{U} \otimes d\mathbf{U}^* + d\mathbf{U} \otimes d\mathbf{U}^* \\ &= \left(j\mathbf{T}dz\mathbf{U} - \frac{\sigma^2}{8N^2}\mathbf{I}dz\mathbf{U}\right) \otimes \mathbf{U}^* + \mathbf{U} \otimes \left(-j\mathbf{T}^*dz\mathbf{U}^* - \frac{\sigma^2}{8N^2}\mathbf{I}dz\mathbf{U}^*\right) \\ &\quad + \langle\mathbf{T}dz \otimes \mathbf{T}^*dz\rangle(\mathbf{U} \otimes \mathbf{U}^*) \\ &= (j\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes j\mathbf{T}^*dz)(\mathbf{U} \otimes \mathbf{U}^*) \\ &\quad + \left(-\frac{\sigma^2}{4N^2}\mathbf{I}dz + \langle\mathbf{T}dz \otimes \mathbf{T}^*dz\rangle\right)(\mathbf{U} \otimes \mathbf{U}^*), \end{aligned} \quad (5.65)$$

where $\otimes$ represents the Kronecker product of matrices, a fundamental operation in matrix algebra. It is noteworthy that the derivation of Eqs. (5.64 and 5.65) relies heavily on a crucial matrix property $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$ [53]. The average value of the first term in the last expression of Eq. (5.65) equals zero. This arises from the fact the average of $\mathbf{T}dz$ is zero, i.e. $\langle\mathbf{T}dz\rangle = 0$. In addition to that, $\mathbf{T}dz$ is the incremental part in the Markovian process, which is independent with respect to $(\mathbf{U} \otimes \mathbf{U}^*)$. Therefore, we derive the ODE for the 2nd order moments of matrix $\mathbf{U}$ as,

$$\begin{aligned} \frac{d\langle\mathbf{U} \otimes \mathbf{U}^*\rangle}{dz} &= \mathbf{S}\langle\mathbf{U} \otimes \mathbf{U}^*\rangle \\ \mathbf{S} &= -\frac{\sigma^2}{4N^2}\mathbf{I} + \frac{\langle\mathbf{T}dz \otimes \mathbf{T}^*dz\rangle}{dz}, \end{aligned} \quad (5.66)$$

$\mathbf{S}$ is made of two primary terms, each with distinct origins and physical interpretations. The first term originates from the derivative of the first-order moments.

Conversely, the second term arises from the property of Brownian motion and signifies the variation resulting from the coupling effect among the second-order moments. Notably, all these terms are weighted by σ^2 . Despite the presence of dz in the expression for $\mathbf{S}$, it ultimately cancels out due to the application of Eq. (5.57a). One may compute the elements of the second term in $\mathbf{S}$ by setting the non-zero σ^2 values solely at positions corresponding to the $((m-1)N+m)$ th row and $((n-1)N+n)$ th column of the matrix, with all other elements being zero.

Utilizing the insights gained from Eq. (5.65) and employing analogous procedures, one can derive the SDE for the fourth-order moments.

$$\begin{aligned}
 & d((\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*)) \\
 &= d(\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*) + (\mathbf{U} \otimes \mathbf{U}^*) \otimes d(\mathbf{U} \otimes \mathbf{U}^*) + d(\mathbf{U} \otimes \mathbf{U}^*) \otimes d(\mathbf{U} \otimes \mathbf{U}^*) \\
 &= ((j\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes j\mathbf{T}^*dz + Sdz)(\mathbf{U} \otimes \mathbf{U}^*)) \otimes (\mathbf{U} \otimes \mathbf{U}^*) \\
 &\quad + (\mathbf{U} \otimes \mathbf{U}^*) \otimes ((j\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes j\mathbf{T}^*dz + Sdz)(\mathbf{U} \otimes \mathbf{U}^*)) \\
 &\quad - ((\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz))((\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*)),
 \end{aligned} \tag{5.67}$$

where the following formula has been implemented

$$\begin{aligned}
 & d(\mathbf{U} \otimes \mathbf{U}^*) \otimes d(\mathbf{U} \otimes \mathbf{U}^*) \\
 &= -((\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz))((\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*)).
 \end{aligned} \tag{5.68}$$

Taking the average on both sides of Eq. (5.57a), we have

$$\begin{aligned}
 & \frac{d\langle (\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*) \rangle}{dz} = \mathbf{F} \langle (\mathbf{U} \otimes \mathbf{U}^*) \otimes (\mathbf{U} \otimes \mathbf{U}^*) \rangle \\
 & \mathbf{F} = \mathbf{S} \otimes \mathbf{I} + \mathbf{I} \otimes \mathbf{S} - \frac{\langle (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \rangle}{dz}.
 \end{aligned} \tag{5.69}$$

In the context of the matrix $\mathbf{F}$, the first and second terms originate from the derivative of the second-order moments. Meanwhile, the third-order term emerges as a consequence of the property of Brownian motion, representing the variation which is induced by the coupling effect among the fourth-order moments. All of these terms are weighted by σ^2 . Notably, the third term of matrix $\mathbf{F}$ can be partitioned into four distinct matrices, namely $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$.

$$\begin{aligned}
& \frac{\langle (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{T}dz \otimes \mathbf{I} - \mathbf{I} \otimes \mathbf{T}^*dz) \rangle}{dz} \\
&= \mathbf{A} + \mathbf{B} - \mathbf{C} - \mathbf{D} \\
&\mathbf{A} = \frac{\langle (\mathbf{T}dz \otimes \mathbf{I}) \otimes (\mathbf{T}dz \otimes \mathbf{I}) \rangle}{dz} \\
&\mathbf{B} = \frac{\langle (\mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{I} \otimes \mathbf{T}^*dz) \rangle}{dz} \\
&\mathbf{C} = \frac{\langle (\mathbf{T}dz \otimes \mathbf{I}) \otimes (\mathbf{I} \otimes \mathbf{T}^*dz) \rangle}{dz} \\
&\mathbf{D} = \frac{\langle (\mathbf{I} \otimes \mathbf{T}^*dz) \otimes (\mathbf{T}dz \otimes \mathbf{I}) \rangle}{dz}.
\end{aligned} \tag{5.70}$$

$\mathbf{A}(i,j)$, $\mathbf{B}(i,j)$, $\mathbf{C}(i,j)$, $\mathbf{D}(i,j)$ can be computed by the following formulas:

$$\begin{aligned}
&\mathbf{A} \left((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k, (m'-1)(2N)^3 + (n'-1)(2N)^2 + (l'-1)2N + k' \right) \\
&= \begin{cases} \sigma^2 & m=l', l=m', n=n', k=k' \\ 0 & \text{otherwise} \end{cases} \\
&\mathbf{B} \left((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k, (m'-1)(2N)^3 + (n'-1)(2N)^2 + (l'-1)2N + k' \right) \\
&= \begin{cases} \sigma^2 & m=m', l=l', n=k', k=n' \\ 0 & \text{otherwise} \end{cases} \\
&\mathbf{C} \left((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k, (m'-1)(2N)^3 + (n'-1)(2N)^2 + (l'-1)2N + k' \right) \\
&= \begin{cases} \sigma^2 & n=n', l=l', m=k, m'=k' \\ 0 & \text{otherwise} \end{cases} \\
&\mathbf{D} \left((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k, (m'-1)(2N)^3 + (n'-1)(2N)^2 + (l'-1)2N + k' \right) \\
&= \begin{cases} \sigma^2 & m=m', k=k', n=l, n'=l' \\ 0 & \text{otherwise} \end{cases},
\end{aligned} \tag{5.71}$$

where the indices m, n, l, k and their primed counterparts m', n', l', k' represent integer numbers ranging from 1 to $2N$. The derivation of the expressions for the elements within matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ relies crucially on Eq. (5.57a) and the properties of the Kronecker product. Equations (5.66 and 5.69) are ODEs with analytical solution:

$$\begin{aligned}
&\langle \mathbf{U}(z) \otimes \mathbf{U}(z)^* \rangle = \exp(\mathbf{S}z) \langle \mathbf{U}(0) \otimes \mathbf{U}(0)^* \rangle \\
&\langle (\mathbf{U}(z) \otimes \mathbf{U}(z)^*) \otimes (\mathbf{U}(z) \otimes \mathbf{U}(z)^*) \rangle \\
&= \exp(\mathbf{F}z) \langle (\mathbf{U}(0) \otimes \mathbf{U}(0)^*) \otimes (\mathbf{U}(0) \otimes \mathbf{U}(0)^*) \rangle,
\end{aligned} \tag{5.72}$$

where $\exp$ denotes the matrix exponential function, a mathematical operation that can be performed analytically by leveraging the properties of the Jordan canonical form. This approach allows for a precise and efficient computation of the matrix exponential, which is often necessary in the analysis of dynamical systems. At the initial condition $z = 0$, it is assumed that the transfer matrix $\mathbf{T}$ is the identity matrix,

reflecting the absence of any evolution of the system at this point. Consequently, the matrices appearing in Eq. (5.72) are also expected to be identity matrices at $z = 0$.

5.2.3 Asymptotic Behaviors of the Matrix Element Moments

5.2.3.1 The 2nd Order Moments

By analyzing Eq. (5.66), it can be inferred that the eigen values of matrix $\mathbf{S}$ are:

$$\frac{\sigma^2}{8N^3}(-2N, -2N \dots, -2N, 0), \quad (5.73)$$

Matrix $\mathbf{S}$ can be expressed as the sum of its eigenvalues, each multiplied by its corresponding eigenvector and the eigenvector's Hermitian transpose. From Eq. (5.73), it is evident that one of the eigenvalues is zero, and this specific eigenvalue and its associated eigenvector play a pivotal role in characterizing the asymptotic behavior during the computation of moments. This is due to the fact that Eq. (5.72) indicates the evolution of second-order moments which is governed by the matrix exponential $\exp(\mathbf{S}z)$. Notably, the eigenvalues of $\exp(\mathbf{S}z)$ are essentially the exponentials of the eigenvalues given in Eq. (5.73). Consequently, as the propagation distance increases, most eigenvalues of the matrix $\exp(\mathbf{S}z)$ tend towards zero, with the exception of the last one, which approaches 1 (i.e., $\exp(0z)$). Consequently, one has:

$$\begin{aligned} & \lim_{z \rightarrow \infty} \langle \mathbf{U}(z) \otimes \mathbf{U}(z)^* \rangle \\ &= \lim_{z \rightarrow \infty} \exp(\mathbf{S}z) \\ &= \lim_{z \rightarrow \infty} \left(\sum_{l=1}^{4N^2-1} \mathbf{v}_l \exp\left(-\frac{\sigma^2}{4N^2}z\right) \mathbf{v}_l^H + \mathbf{v} \exp(0z) \mathbf{v}^H \right) \\ &= \mathbf{v} \mathbf{v}^H, \end{aligned} \quad (5.74)$$

where $\mathbf{v}$ is the eigenvector corresponding to the eigenvalue 0 of matrix $\mathbf{S}$, and $\mathbf{v}_l$ represents the eigenvectors associated with the remaining negative eigenvalues of $\mathbf{S}$. We leverage the fundamental property of the matrix exponential, which states that the eigenvectors of matrix $\mathbf{S}$ remain as eigenvectors of the matrix exponential $\exp(\mathbf{S}z)$ [54]. Hence, $\mathbf{v}$ can be expressed by

$$\mathbf{v}(k) = \begin{cases} \frac{1}{\sqrt{2N}} & k = (m-1)2N + m \\ 0 & \text{otherwise} \end{cases}, \quad (5.75)$$

where m is an integer, $m = 1, 2, \dots, 2N$. Based on Eqs. (5.74 and 5.75), one may deduce the following conclusions:

$$\lim_{z \rightarrow \infty} \langle |U_{ij}|^2 \rangle = \frac{1}{2N}, \quad (5.76)$$

which is in exact agreement with the Haar matrix property stated in [43].

5.2.3.2 The 4th Order Moments

Analogous to the preceding procedures, the matrix $\mathbf{F}$ in Eq. (5.69) possesses numerous negative eigenvalues along with two eigenvalues that are zero. Therefore, one has:

$$\begin{aligned} & \lim_{z \rightarrow \infty} \langle (\mathbf{U}(z) \otimes \mathbf{U}(z)^*) \otimes (\mathbf{U}(z) \otimes \mathbf{U}(z)^*) \rangle \\ &= \lim_{z \rightarrow \infty} \exp(\mathbf{F}z) \\ &= \mathbf{v}_1 \exp(0z) \mathbf{v}_1^H + \mathbf{v}_2 \exp(0z) \mathbf{v}_2^H \\ &= \mathbf{v}_1 \mathbf{v}_1^H + \mathbf{v}_2 \mathbf{v}_2^H, \end{aligned} \quad (5.77)$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are the two eigen vectors specifically related to two zero eigenvalues of $\mathbf{F}$. The elements of the vectors should be formulated as

$$\begin{aligned} & \mathbf{v}_1((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k) \\ &= \begin{cases} \frac{1}{\sqrt{2(2N-1)2N}} & m = n, l = k, m \neq l \\ -\frac{1}{\sqrt{2(2N-1)2N}} & m = k, n = l, m \neq n \\ 0 & \text{otherwise} \end{cases} \\ & \mathbf{v}_2((m-1)(2N)^3 + (n-1)(2N)^2 + (l-1)2N + k) \\ &= \begin{cases} \frac{1}{\sqrt{(2N+1)N}} & m = n = l = k \\ \frac{1}{\sqrt{(2N+1)4N}} & m = n, l = k, m \neq l; \\ & m = k, n = l, m \neq n \\ 0 & \text{otherwise} \end{cases} \end{aligned} \quad (5.78)$$

where m, n, l, k are integers ranging from 1 to $2N$. It is inferred from Eqs. (5.77 and 5.78) that the asymptotic behaviors of the fourth-order matrix element moments can be derived as detailed below.

$$\begin{aligned}
\lim_{z \rightarrow \infty} \langle |U_{ij}|^4 \rangle &= \frac{2}{2N(2N+1)} \\
\lim_{z \rightarrow \infty} \langle |U_{ik}|^2 |U_{il}|^2 \rangle &= \frac{1}{2N(2N+1)} \quad (k \neq l) \\
\lim_{z \rightarrow \infty} \langle |U_{ik}|^2 |U_{jl}|^2 \rangle &= \frac{1}{(2N-1)(2N+1)} \quad (i \neq j, k \neq l) \\
\lim_{z \rightarrow \infty} \langle U_{ik} U_{jl} U_{il}^* U_{jk}^* \rangle &= \frac{-1}{(2N-1)2N(2N+1)} \quad (i \neq j, k \neq l),
\end{aligned} \tag{5.79}$$

which is in precise accordance with the Haar matrix property as described in [30, 43].

To investigate the Kerr effect inside multimode fibers, we consider the nonlinear term on the k th mode $\langle \hat{N}_k \rangle$ [6]:

$$\langle \hat{N}_k \rangle = j\gamma \sum_{l=1}^N \sum_{m=1}^N \sum_{n=1}^N \frac{2(\langle |U_{ml}|^2 |U_{nk}|^2 \rangle + \langle U_{ml} U_{nk} U_{mk}^* U_{nl}^* \rangle)}{(1+d_{kl})(1+d_{mn})} f_{mmnn} |A_l|^2 A_k, \tag{5.80}$$

where N is the number of the modes in the multimode fiber, A_l denotes the l th mode's amplitude, f_{mmnn} stands for the nonlinear overlap factor calculated from the mode distribution and index distribution, γ is the fiber nonlinear coefficient related to the mode field and the Kerr coefficient, d_{kl} is a number, which is defined as 1 if mode k and mode l belong to the same mode group and 0 otherwise.

Equation (5.80) indicates that the derived formulas are readily adaptable for direct implementation in the context of evolutionary analysis of the Kerr effect.

5.2.4 Results and Discussions

Monte Carlo simulations have been executed to validate the proposed analytical expressions for calculating the evolution of transfer matrix element moments. These simulations were performed numerically using Eq. (5.54), wherein the vector $\vec{\beta}'$ was modeled as a $4N^2$ -dimensional Brownian motion with equal variance across all components and no correlation among them [39]. A total of 3000 realizations were computed to determine the average values for the second-order and fourth-order moments of the matrix elements. The propagation distance was normalized to $\sigma^2 z / (4N^2)$, with z representing the fiber length.

The analytical results were directly calculated using Eq. (5.72). In the numerical simulations, the normalized distance was divided into steps of 0.01. In terms of computational time, the analytical method exhibits a comparable cost to that of a single realization in the numerical simulations. Consequently, by implementing the analytical expressions, the overall computational cost was reduced by three orders of magnitude.

Two multimode systems were investigated: a 2-mode system consisting of one LP mode with two polarized sub-modes, and a 6-mode system comprising three LP

modes with six polarized sub-modes. These systems will be referred to as the 2-mode and 6-mode systems, respectively, throughout the remainder of this section.

Initially, Fig. 5.8 presents the outcomes for the second-order moments in the 2-mode system. The impeccable alignment between the analytical predictions and Monte Carlo simulations underscores the veracity of the proposed analytical formulas. The relative differences between the two results for various moments during their evolution, as depicted in Fig. 5.9, reveal discrepancies of less than 3%. Given the unitary nature of the matrix, one has $|U_{11}|^2 = |U_{22}|^2$, $|U_{12}|^2 = |U_{21}|^2$. As the number of realizations in the Monte Carlo simulations increases, these discrepancies can be further minimized. Throughout the remainder of the analysis, the discrepancies between numerical and analytical results are comparable to those presented in Fig. 5.9. In Fig. 5.8, it is evident that the second-order moments of the transfer matrix elements gradually converge towards the value of 1/2, consonant with the asymptotic predictions. The evolutionary trends of the curves suggest that the convergence is nearly complete when $\sigma^2 z/(4N^2)$ reaches approximately 4, which is a relatively small value. According to [39], the mean square of the modal dispersion vector $\vec{\tau}$ is proportional to the fiber length [39, 42]:

$$\langle \tau^2 \rangle = \gamma^2 z = \frac{\sigma^2 z}{\omega^2 4N^2} (4N^2 - 1), \quad (5.81)$$

To approximately determine the fiber length necessary for achieving the condition of $\sigma^2 z/(4N^2) = 4$, let's select a zero reference frequency as our starting point.

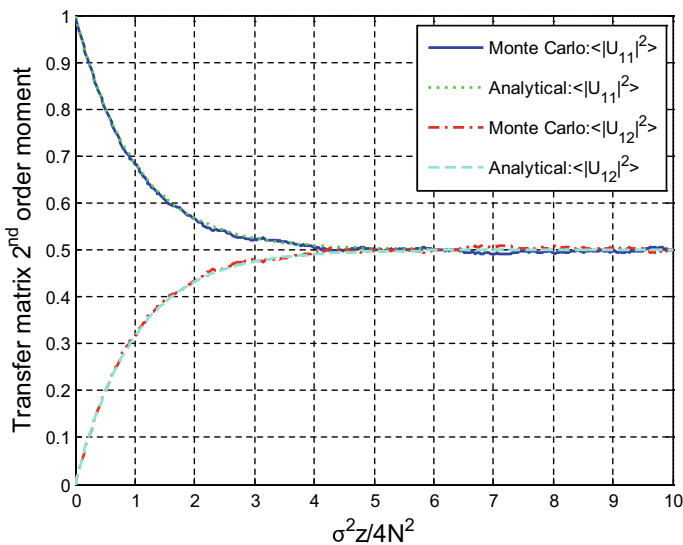


Fig. 5.8 The evolution of the 2nd order transfer matrix element moments for the 2-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

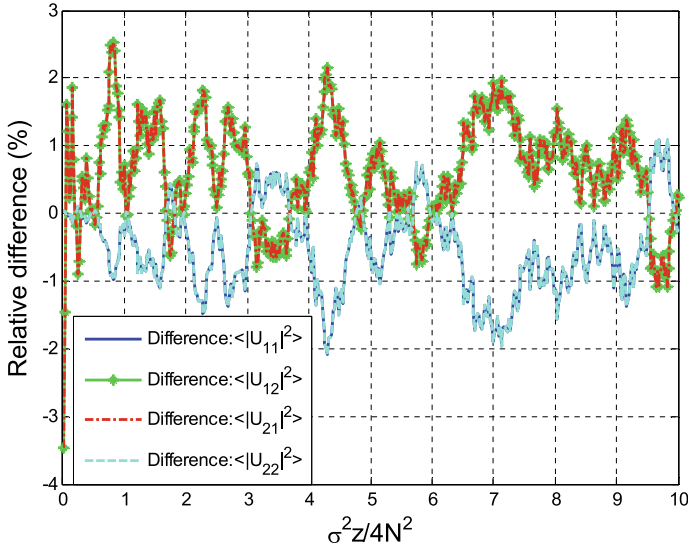


Fig. 5.9 The relative differences between the analytical and the numerical results for the various moments during the evolution. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

Assuming a signal wavelength of 1550 nm, we consider the mean-square magnitude of the generalized birefringence parameter, γ as $1 \text{ ps}/\sqrt{\text{km}}$. Given that the number of modes in the system, N , is 3, we can calculate: $\sigma^2/(4N^2) = 42.2/\text{m}$, from which we derive $\sigma^2 = 1521/\text{m}$. The condition $\sigma^2 z/(4N^2) = 4$ implies that the fiber length, z , should be approximately 0.1 m, when $\sqrt{\langle\tau^2\rangle}$ is about 0.01 ps. If γ has the value of $0.1 \text{ ps}/\sqrt{\text{km}}$, $\sigma^2/(4N^2) = 0.422/\text{m}$, $\sigma^2 = 15.21/\text{m}$ and the corresponding z will be about 10 m when $\sqrt{\langle\tau^2\rangle}$ is about 0.01 ps.

When investigating the inter-channel nonlinearity, specifically the inter-channel cross-phase modulation effect, it is appropriate to select the center frequency of a neighboring channel as the reference frequency, and assign the channel angular frequency spacing to ω . For instance, if the channel frequency spacing is chosen as 100 GHz, we proceed with the analysis. Assuming the mean-square magnitude of the generalized birefringence parameter, γ , has a value of $1 \text{ ps}/\sqrt{\text{km}}$, and the number of modes in the system, N , is 3, we can calculate the following: $\sigma^2/(4N^2) = 1.128 \times 10^{-5}/\text{m}$, which implies that $\sigma^2 = 4.06 \times 10^{-4}/\text{m}$. $\sigma^2 z/(4N^2) = 4$ suggests that z is about 10 km, when $\sqrt{\langle\tau^2\rangle}$ is about 3.2 ps. If γ has the value of $0.1 \text{ ps}/\sqrt{\text{km}}$, $\sigma^2/(4N^2) = 1.128 \times 10^{-7}/\text{m}$, $\sigma^2 = 4.06 \times 10^{-6}/\text{m}$ and the corresponding z will be about 1000 km when $\sqrt{\langle\tau^2\rangle}$ is about 3.2 ps.

For a given reference signal center frequency, an increase in the $\sqrt{\langle\tau^2\rangle}$ can result in a shorter fiber length L . In parameter amplifiers employing high nonlinear fibers

(HNLFs) or in multimode sensing systems, fibers typically have tens to hundreds of meters in length. Conversely, in data center transmission links, fiber lengths are on the order of just a few meters. As a consequence, the evolution of second-order moments (which characterize statistical fluctuations) becomes negligible in long-haul fiber transmission systems extending hundreds of kilometers. However, in the aforementioned practical applications, these fluctuations are significant and can induce power/phase variations due to variations in the transfer matrix elements. This underscores the importance for designers to consider enhanced dynamic range capabilities and implement more robust fault tolerance mechanisms to ensure reliable system performance.

Subsequently, Figs. 5.10 and 5.11 present the fourth-order matrix element moments for the previously discussed two-mode system. The four crucial representative fourth-order moments, as mentioned in Eq. (5.79), are selected and depicted in these figures. Once again, a perfect alignment is observed between the analytical predictions and the Monte Carlo simulation results. Analogous to the trends observed in Fig. 5.8, the convergence of the curves is achieved within the normalized propagation distance range of $[0, 4]$ for $\sigma^2 z/(4N^2)$, further emphasizing the significance of moment evolution in short fiber systems. For clarity and convenience of illustration, the imaginary part of the dash-dot line in Fig. 5.11, which is negligible, has been omitted. The asymptotic values of the fourth-order moments are in agreement with Eq. (5.79) and coincide with the findings reported in the published literature, validating the analysis and simulations presented in [30, 47].

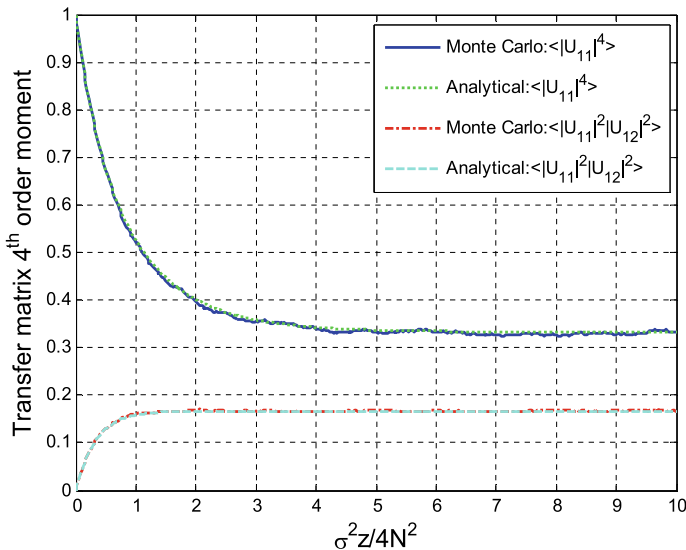


Fig. 5.10 The evolution of the 4th order transfer matrix element moments for the 2-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

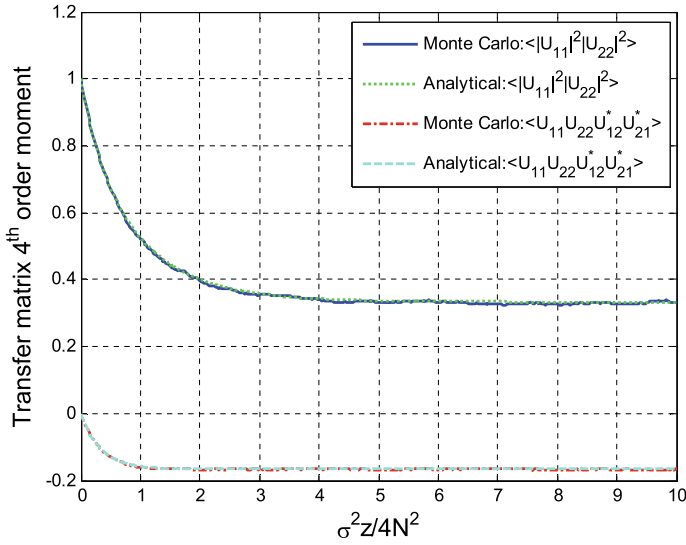


Fig. 5.11 The evolution of the 4th order transfer matrix element moments for the 2-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

Having validated the formulas for the two-mode system, we now present the outcomes for the six-mode system in Figs. 5.12, 5.13, and 5.14. Figure 5.12 illustrates the evolutionary behavior of the second-order moments for the six-mode system, demonstrating the anticipated perfect alignment between analytical results and Monte Carlo simulations. The convergence process is achieved within the range of $[0, 2.5]$ for the normalized propagation distance $\sigma^2 z / (4N^2)$. The converged values of the second-order moments coincide with the predicted asymptotic values of $1/6$. Figures 5.13 and 5.14 show the results of the fourth-order matrix element moments for the six-mode system, once again revealing a seamless match between analytical predictions and Monte Carlo averaged values. Similar to the two-mode system, the four pivotal representative fourth-order moments are depicted separately in these two figures. The convergence behavior mirrors that of the second-order moments, accomplished within the same range of $[0, 2.5]$ for $\sigma^2 z / (4N^2)$. The asymptotic values of the moments align with published findings. Notably, a portion of the fourth-order (dashed) line in Fig. 5.14 has been magnified to facilitate a closer examination of the detailed values within the convergence regime.

5.2.5 Summary

A novel analytical framework is introduced to investigate the evolutionary patterns of transfer matrix elements in multimode systems. At its core, a fundamental SDE

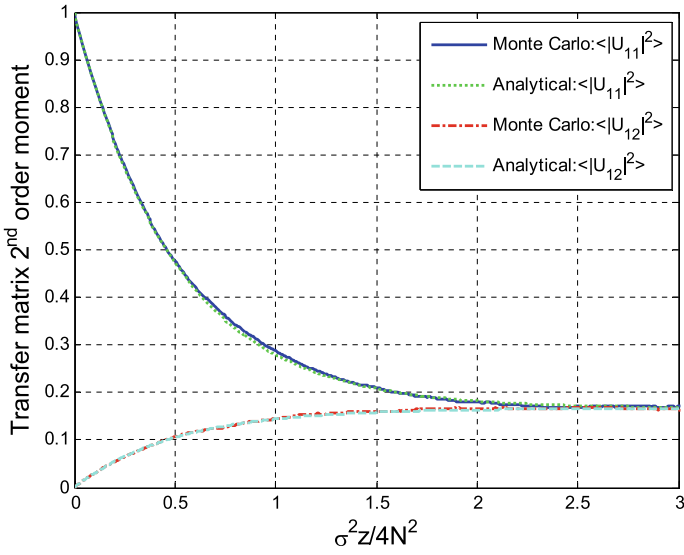


Fig. 5.12 The evolution of the 2nd order transfer matrix element moments for the 6-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

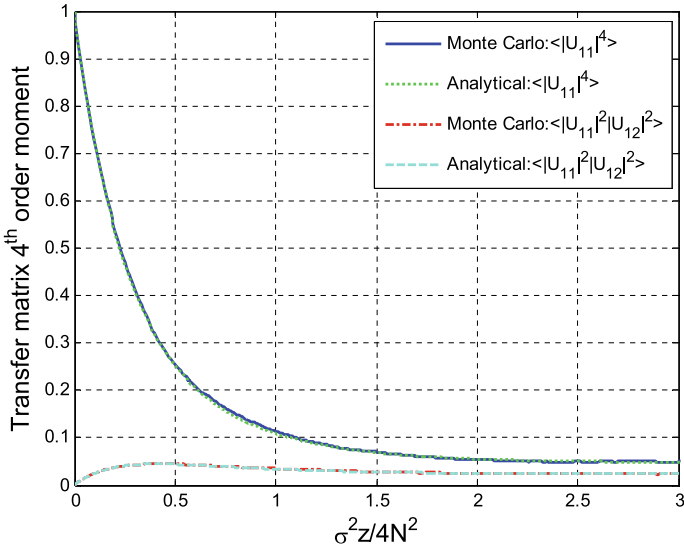


Fig. 5.13 The evolution of the 4th order transfer matrix element moments for the 6-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

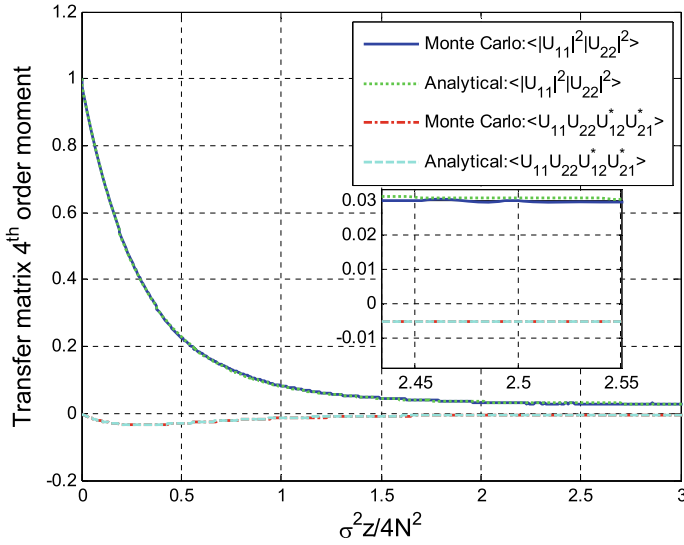


Fig. 5.14 The evolution of the 4th order transfer matrix element moments for the 6-mode system. This figure is reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24]

is derived, constituting a foundational pillar for subsequent derivations of higher-order moment equations. By averaging these SDEs, ODEs are formulated, and their analytical solutions are obtained. A thorough examination of these ODEs reveals that the asymptotic behaviors of their solutions are in excellent agreement with previously published results, which were derived via an entirely distinct methodology.

To further substantiate the validity of these analytical solutions, a comparison is made with Monte Carlo simulation outcomes, yielding a flawless concordance. This proposed methodology holds significant potential for facilitating the design and analysis of future multimode fiber links or multimode fiber sensing systems, offering a robust analytical tool to predict and understand their behavior.

5.3 Analytical Expressions for the Crosstalk of Super-Modes in the Tightly Bounded Multicore Fibers

Part of the section reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55].

In this section, we extend the application of the random coupled mode theory from multimode fibers to multicore fibers. An analytical framework is devised specifically for the super-modes in tightly confined multicore fibers, which incorporates both

deterministic and random inter-core coupling effects. This analysis is grounded in the derivation of ordinary differential equations (ODEs) from their stochastic counterparts (SDEs). Theoretically, it is uncovered that the crosstalk level scales directly with the square of the ratio between random and deterministic inter-core coupling strengths, and inversely with the random coupling correlation length.

To further enrich our understanding of these tightly bound multicore fibers, ODEs for the super-mode power variance and power correlations are also presented. In the scenario of weak super-mode crosstalk, concise and straightforward formulas are provided for estimating both the crosstalk power and power covariance among super-modes. These analytical tools offer an efficient approach to evaluate and analyze the performance of multicore fibers in practical applications.

5.3.1 Introduction

As is shown in Chap. 3, multicore fibers (MCFs) are of great importance with various applications including optical fiber communications [25, 56–58]. Transmission systems utilizing MCFs can be categorized into two primary types: those based on digital signal processing (DSP) and those that are non-DSP-based. The DSP-based systems employ DSP chips to counteract inter-core or inter-mode crosstalk, with the primary objective of mitigating delay spread [39–42, 50, 59]. Conversely, non-DSP-based systems focus on minimizing the crosstalk occurring between different transmission modes.

There are two primary design approaches for minimizing crosstalk in multicore fibers (MCFs). The first approach targets reducing inter-core crosstalk by utilizing core trenches, implementing a large refractive index difference between cores, and maintaining a relatively long core pitch (the distance between adjacent core centers) [60–71]. This design helps isolate the cores from each other. The second approach, on the other hand, employs super-modes for transmission and aims to minimize inter-super-mode crosstalk (ISMXT) by strengthening the coupling between cores. This results in a more significant difference in the propagation constants of the super-modes, allowing for a much shorter core pitch [72–75]. In some cases, the core pitch is even reduced to twice the core radius, effectively eliminating spacing between adjacent cores [72]. Compared to the first approach, this design enables a closer arrangement of cores, enabling MCFs to accommodate a higher number of cores/transmission modes within a given cross-sectional area.

Crosstalk analysis holds paramount significance in the MCF design. The random crosstalk occurring between cores [60–71] has gained extensive attention, with investigations mostly based on the coupled mode equations (CMEs). Multiple models have emerged, aiming to tackle this challenge with enhanced accuracy and efficiency. Notably, in [71], the polarization mode dispersion (PMD) was included in the coupling equation, revealing that when the birefringence correlation length aligned closely with the coupling correlation length, PMD tended to mitigate crosstalk.

Despite the groundbreaking efforts in [60–71], there has been a notable scarcity of theoretical investigations devoting the random crosstalk phenomena among super-modes. Saitoh et al. [72] experimentally measured crosstalk among super-modes in a four-core fiber. Xia et al. [73, 74] by means of numerical simulations, designed a three-core fiber and proposed enhancing the effective index difference among super-modes as a means to mitigate crosstalk. However, these studies fell short of providing quantitative theoretical predictions for super-mode crosstalk [73, 74]. In [75], while deterministic crosstalk among OAM super-modes was discussed, the random counterpart remained unexplored. This analytical gap is the result of the fundamental distinction between super-mode transmission and conventional multi-core fiber (MCF) transmission within individual cores [72–76]. Super-mode transmission more closely resembles multimode transmission in MMFs, rendering traditional crosstalk evaluation methods between cores [60, 71] inapplicable to super-mode crosstalk analysis. The chief challenge in formulating a comprehensive theory for random crosstalk among super-modes lies in the interplay between deterministic and random coupling effects. Their comparable coupling beat lengths and correlation lengths significantly complicate the theoretical derivation process.

In our recent endeavor on crosstalk in multi-core fibers (MCFs) [77], we introduced a unitary transformation approach to investigate random crosstalk within MCFs featuring deterministic coupling. However, the MCFs analyzed in [77] exhibited a deterministic coupling beat length that was comparable to the polarization mode dispersion (PMD) correlation length, indicating that the deterministic coupling strength was insufficient to maintain stable super-modes. Furthermore, [77] did not provide the explicit expressions that describe the crosstalk behavior among super-modes in detail.

In this work, we present a comprehensive analytical framework tailored specifically for the random crosstalk analysis of super-modes in tightly coupled MCFs. This theory is rooted in the application of SDEs. By leveraging a unitary transformation, we effectively eliminate deterministic coupling among cores while simultaneously redistributing inter-core random coupling effects across super-modes. In MCFs designed specifically for super-mode transmission, the deterministic coupling strength is inherently strong, leading to a deterministic coupling beat length significantly shorter than the birefringence correlation length. Consequently, the influence of polarization effects on the ISMXT becomes negligible. Employing the transformed SDE tailored for super-modes, we derive coupled ordinary differential equations (ODEs) which govern super-mode power, its variance, and power correlations. These ODEs are invaluable for estimating system crosstalk budgets and analyzing system capacity. Furthermore, under the assumption of weak ISMXT, we provide concise and straightforward formulas that characterize crosstalk among super-modes, facilitating a deeper understanding of this intricate phenomenon.

5.3.2 Theory

In the context of tightly bound MCFs specifically engineered for super-mode transmission, the deterministic coupling beat length, which measures in millimeters [72], is significantly shorter than the polarization mode dispersion (PMD) correlation length, typically in the order of meters [71]. This enables us to disregard the influence of PMD on super-mode crosstalk, as the signal polarization remains virtually unchanged within the coupling beat length. To validate this approximation, a Monte Carlo simulation will be conducted which incorporates PMD effects, and the outcomes will be benchmarked against the analytical results presented here. Consequently, the scalar coupled mode theory (CMT) without considering the polarization effect serves as an appropriate tool to model the propagation dynamics of amplitudes within the cores of these MCFs.

$$\begin{aligned} \frac{d\mathbf{E}}{dz} &= i\mathbf{B}\mathbf{E} + i\boldsymbol{\kappa} \circ \mathbf{f}(z)\mathbf{E} \\ \mathbf{B} &= \begin{pmatrix} \beta_1 & \beta_{12} & & \beta_{1N} \\ & \beta_2 & & \\ & & \ddots & \\ \beta_{N1} & & \beta_{NN-1} & \beta_N \end{pmatrix}, \\ \boldsymbol{\kappa} \circ \mathbf{f}(z) &= \begin{pmatrix} \kappa_{11}f_{11}(z) & \cdots & \kappa_{1N}f_{1N}(z) \\ \vdots & \kappa_{22}f_{22}(z) & \vdots \\ & & \ddots & \\ \kappa_{N1}f_{N1}(z) & \cdots & \kappa_{NN}f_{NN}(z) \end{pmatrix}, \end{aligned} \quad (5.82)$$

where $\mathbf{E}$ represents a vector comprising the amplitudes E_m of each core, where m denotes the specific core. The deterministic coupling matrix $\mathbf{B}$ encapsulates the propagation characteristics of these cores, with its diagonal elements β_m representing the propagation constants of the m th core, which incorporate chromatic dispersion effects. The non-diagonal elements β_{mn} within $\mathbf{B}$ signify the deterministic coupling coefficients between the m th and n th cores. In addition to the deterministic coupling, there exists a random coupling component incorporated in the matrix $\boldsymbol{\kappa} \circ \mathbf{f}(z)$, where κ_{mn} and f_{mn} represent the random coupling amplitude and the z -directional random variation process, respectively, between the m th and n th cores. This random coupling arises from perturbations such as fiber bending and core misalignment. The random variation process $f_{mn}(z)$ is modeled as a filtered color noise, characterized by a specific correlation function that describes its statistical properties over the fiber length [61, 62, 70, 71]

$$\langle f_{mn}(z + z') f_{mn}^*(z') \rangle = \eta_{mn}(z). \quad (5.83)$$

While it is not an absolute requirement, it is commonly assumed that the random variation processes f_{mn} are identically distributed and independent (i.i.d.) variables. This assumption simplifies the analysis by treating each f_{mn} as statistically equivalent and uncorrelated with others. Furthermore, the correlation function η_{mn} of these processes is often modeled as a two-sided exponential function [61, 62, 70, 71]:

$$\eta_{mn}(z) = \eta(z) = \exp\left(-\frac{|z|}{l_c}\right). \quad (5.84)$$

where l_c is the correlation length. η has its Fourier transform as

$$\mathbf{FT}(\eta)(\omega) = \frac{2l_c}{1 + l_c^2 \omega^2}, \quad (5.85)$$

where $\mathbf{FT}$ denotes the Fourier transform, and ω represents the angular frequency.

Equation (5.82) was solved in case $\mathbf{B}$ was diagonal [61, 62, 70]. A more general case when it is not diagonal is studied here. Because $\mathbf{B}$ is Hermitian, matrix decomposition can be applied as $\mathbf{B} = \mathbf{H}\mathbf{D}\mathbf{H}^H$, with $\mathbf{D}$ to be diagonal and $\mathbf{H}$ to be unitary [77]. Therefore, one may transform Eq. (5.82) to Eq. (5.86) [77].

$$\begin{aligned} \mathbf{E}' &= \mathbf{H}^H \mathbf{E}, \\ \kappa'(z) &= \mathbf{H}^H \kappa \circ \mathbf{f}(z) \mathbf{H}, \\ \frac{d\mathbf{E}'}{dz} &= i\mathbf{D}\mathbf{E}' + i\kappa'(z)\mathbf{E}'. \end{aligned} \quad (5.86)$$

We may compute the correlation function of the matrix $\kappa'(z)$ element as

$$\langle \kappa'_{mn}(z + z') \kappa'^*_{mn}(z') \rangle = \left(\sum_{k=1}^N \sum_{l=1}^N |H_{km}|^2 |H_{ln}|^2 |\kappa_{kl}|^2 \right) \eta(z). \quad (5.87)$$

In MCFs specifically engineered for super-mode transmission, the super-modes are non-degenerate and possess distinct propagation constants, resulting in a diagonal matrix $\mathbf{D}$ with unique diagonal elements ($d_1, \dots, d_N$). By substituting $\mathbf{A} = \exp(-i\mathbf{D}z)\mathbf{E}'$, we may recast Eq. (5.86) into [61]

$$\frac{d\mathbf{A}}{dz} = \exp(-i\mathbf{D}z) i\kappa'(z) \exp(i\mathbf{D}z) \mathbf{A}. \quad (5.88)$$

One may define matrix $\mathbf{F}$ as

$$\mathbf{F} = \exp(-i\mathbf{D}z) i\kappa'(z) \exp(i\mathbf{D}z). \quad (5.89)$$

If the correlation length l_c is significantly shorter compared to the propagation length, the elements of the matrix $\mathbf{F}$ exhibit the following property:

$$\begin{aligned}
\langle F_{mn} dz F_{mn}^* dz \rangle &= K_{mn} dz, \\
\langle F_{mn} dz F_{m'n'}^* dz \rangle &= 0 (m \neq m', n \neq n'), \\
K_{mn} &= \left(\sum_{k=1}^N \sum_{l=1}^N |H_{km}|^2 |H_{ln}|^2 |\kappa_{kl}|^2 \right) \mathbf{FT}(\eta)(d_m - d_n),
\end{aligned} \tag{5.90}$$

Equation (5.88) is in the Stratonovich sense, which ought to be converted to Ito sense [77]:

$$d\mathbf{A} = i\mathbf{F}d\mathbf{z}\mathbf{A} - \frac{1}{2} \begin{pmatrix} \sum_{m=1}^N K_{1m} \cdots 0 \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots 0 \cdots & \sum_{m=1}^N K_{Nm} \end{pmatrix} dz\mathbf{A}. \tag{5.91}$$

Defining the power of the m th super-mode as the expectation value of $\langle P_m \rangle = \langle e_m' e_m'^* \rangle = \langle a_m a_m^* \rangle$, where e_m' and a_m represent the elements of the vectors $\mathbf{E}'$ and $\mathbf{A}$ respectively, one can subsequently derive the coupled power equation based on Eqs. (5.90 and 5.91) [3, 13, 77]:

$$\frac{d\langle P_m \rangle}{dz} = \sum_{m'=1}^N K_{mm'} (\langle P_{m'} \rangle - \langle P_m \rangle), \tag{5.92}$$

and the coupled power correlation equation is formulated as well [3, 13, 77]:

$$\frac{d\langle P_m^2 \rangle}{dz} = -2 \sum_{m'=1}^N K_{mm'} \langle P_m^2 \rangle + 4 \sum_{m'=1}^N K_{mm'} \langle P_m P_{m'} \rangle - 2\langle P_m(z)^2 \rangle K_{mm}, \tag{5.93a}$$

$$\begin{aligned}
\frac{d\langle P_m P_n \rangle}{dz} &= \left(- \sum_{m'=1}^N K_{mm'} - \sum_{n'=1}^N K_{nn'} \right) \langle P_m P_n \rangle + \sum_{m'=1}^N K_{mm'} \langle P_n P_{m'} \rangle \\
&\quad + \sum_{n'=1}^N K_{nn'} \langle P_m P_{n'} \rangle - 2K_{mn} \langle P_m P_n \rangle.
\end{aligned} \tag{5.93b}$$

Equations (5.92–5.93b) can be solved analytically [3, 77].

Employing the aforementioned theory, we can analyze a two-core multi-core fiber (MCF). This fiber exhibits a symmetrical super-mode (super-mode 1) and an anti-symmetrical super-mode (super-mode 2). Notably, the inter-super-mode coupling coefficients K_{12} and K_{21} are equal. Therefore, we may solve Eq. (5.92) as

$$\begin{aligned}
\langle P_1(z) \rangle &= \frac{1}{2}(1 + \exp(-2K_{12}z))\langle P_1(0) \rangle + \frac{1}{2}(1 - \exp(-2K_{12}z))\langle P_2(0) \rangle \\
\langle P_2(z) \rangle &= \frac{1}{2}(1 + \exp(-2K_{12}z))\langle P_2(0) \rangle + \frac{1}{2}(1 - \exp(-2K_{12}z))\langle P_1(0) \rangle
\end{aligned} \tag{5.94}$$

One may formulate the ISMXT from super-mode 2 to super-mode 1 as

$$\begin{aligned}
ISMXT &= \frac{1}{2} \frac{(1 - \exp(-2K_{12}z))\langle P_2(0) \rangle}{\langle P_2(0) \rangle} \\
&= \frac{1}{2}(1 - \exp(-2K_{12}z)) \approx K_{12}z (K_{12}z \ll 1)
\end{aligned} \tag{5.95}$$

In scenarios where crosstalk is weak, specifically when $K_{12}z \ll 1$, the ISMXT can be approximated as $K_{12}z$, mirroring the analysis of inter-core crosstalk but with a modified definition of the crosstalk coefficient K_{12} by Eq. (5.87). Analogously, Eq. (5.93a, 5.93b) can be solved as

$$\begin{aligned}
\langle P_1(z)^2 \rangle &= \left(\frac{1}{6} \exp(-6K_{12}z) + \frac{1}{2} \exp(-2K_{12}z) + \frac{1}{3} \right) \langle P_1(0)^2 \rangle \\
&\quad + \left(-\frac{2}{3} \exp(-6K_{12}z) + \frac{2}{3} \right) \langle P_1(0)P_2(0) \rangle \\
&\quad + \left(\frac{1}{6} \exp(-6K_{12}z) - \frac{1}{2} \exp(-2K_{12}z) + \frac{1}{3} \right) \langle P_2(0)^2 \rangle \\
\langle P_1(z)P_2(z) \rangle &= \left(\frac{1}{6} - \frac{1}{6} \exp(-6K_{12}z) \right) \langle P_1(0)^2 \rangle \\
&\quad + \left(\frac{2}{3} \exp(-6K_{12}z) + \frac{1}{3} \right) \langle P_1(0)P_2(0) \rangle \\
&\quad + \left(\frac{1}{6} - \frac{1}{6} \exp(-6K_{12}z) \right) \langle P_2(0)^2 \rangle \\
\langle P_2(z)^2 \rangle &= \left(\frac{1}{6} \exp(-6K_{12}z) + \frac{1}{2} \exp(-2K_{12}z) + \frac{1}{3} \right) \langle P_2(0)^2 \rangle \\
&\quad + \left(-\frac{2}{3} \exp(-6K_{12}z) + \frac{2}{3} \right) \langle P_1(0)P_2(0) \rangle \\
&\quad + \left(\frac{1}{6} \exp(-6K_{12}z) - \frac{1}{2} \exp(-2K_{12}z) + \frac{1}{3} \right) \langle P_1(0)^2 \rangle
\end{aligned} \tag{5.96}$$

Therefore, we may calculate the variance of ISMXT from super-mode 2 to super-mode 1 as

$$\text{var}(ISMXT) = \left(\frac{1}{6} \exp(-6K_{12}z) - \frac{1}{2} \exp(-2K_{12}z) + \frac{1}{3} \right)$$

$$\begin{aligned}
& - \left(\frac{1}{2} (1 - \exp(-2K_{12}z)) \right)^2 \\
& = \frac{1}{12} - \frac{1}{4} \exp(-4K_{12}z) + \frac{1}{6} \exp(-6K_{12}z) \approx (K_{12}z)^2 (K_{12}z \ll 1)
\end{aligned} \tag{5.97}$$

Under the conditions of weak crosstalk, where $K_{12}z \ll 1$, the variance of ISMXT from super-mode 2 to super-mode 1 approximates to $(K_{12}z)^2$. Consequently, the standard deviation is approximately $K_{12}z$. The results presented in Eqs. (5.90, 5.95, and 5.97) can be adapted to evaluate the ISMXT power and power standard deviation between the m th and n th super-modes by substituting the inter-super-mode coupling coefficient K_{12} by K_{mn} if the crosstalk is weak.

As is illustrated, the power coupling coefficient K_{12} , which quantifies the strength of random crosstalk between super-modes, plays a crucial role in characterizing ISMXT. This coefficient can be determined by calculating as follows

$$K_{12} = \frac{2l_c}{1 + l_c^2(2\beta_{12})^2} \left(\frac{1}{4} |\kappa_{12}|^2 + \frac{1}{4} |\kappa_{21}|^2 \right) \tag{5.98}$$

Given that the random coupling amplitude is proportional to the deterministic part by a factor of r , i.e., $|\kappa_{12}| = |\kappa_{21}| = r\beta_{12}$, the limit for K_{12} in scenarios of strong deterministic coupling is proportional to r^2 and inversely proportional to the correlation length l_c :

$$\lim_{\beta_{12} \rightarrow \infty} K_{12} = \frac{r^2}{4l_c} \tag{5.99}$$

5.3.3 Results and Discussions

To validate the proposed analytical formulas, Monte Carlo simulations have been performed. In this context, a two-core fiber, as previously discussed, together with a four-core multi-core fiber (MCF) fabricated in [72] are analyzed in depth. The merit of this four-core fiber lies in its possession of four non-degenerate super-modes, which can facilitate MIMO-less transmission. The cross-sectional views of both the two-core fiber and the four-core fiber are illustrated in Fig. 5.15.

The fiber parameters are consistent with those reported in [72]. The step-index fiber features a cladding index of 1.45, and the core index differs by 0.31% from the cladding index. The core radius measures $2.5 \mu\text{m}$, while the cladding diameter is $125.7 \mu\text{m}$. For the two-core fiber, the core pitch is $8.66 \mu\text{m}$. In the four-core MCF, two distinct core pitches of 5 and $8.66 \mu\text{m}$ are employed to generate non-degenerate super-modes. The signal wavelength is set at $1.55 \mu\text{m}$. For the convenience of readers,

Fig. 5.15 The cross section of **a** the two-core fiber mentioned above. **b** The four-core fiber mentioned in [72]. This figure is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

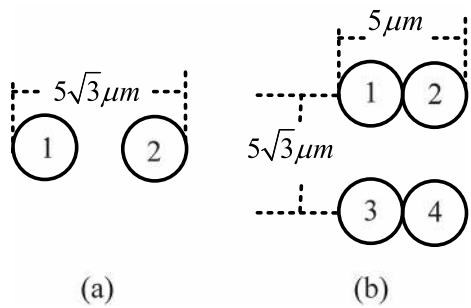


Table 5.3 presents a comprehensive list of detailed parameters for the four-core fiber as described in [72].

In the case of the two-core fiber, the deterministic coupling coefficient is evaluated as $\beta_{12} = 960/\text{m}$. Turning to the four-core fiber, the deterministic coupling coefficients between specific core pairs are: $\beta_{12} = 3031/\text{m}$ between core 1 and core 2, $\beta_{13} = 1285/\text{m}$ between core 1 and core 3, and $\beta_{14} = 960/\text{m}$ between core 1 and core 4. These computed values align well with the reported values in [72]. Notably, while the two-core fiber possesses two super-modes, the four-core MCF hosts four super-modes, each characterized by its unique propagation constant as follows: [72].

$$\begin{aligned}\beta_{LP01-like} &= \beta_0 + \beta_{12} + \beta_{13} + \beta_{14} \\ \beta_{LP11a-like} &= \beta_0 + \beta_{12} - \beta_{13} - \beta_{14} \\ \beta_{LP11b-like} &= \beta_0 - \beta_{12} + \beta_{13} - \beta_{14} \\ \beta_{LP21-like} &= \beta_0 - \beta_{12} - \beta_{13} + \beta_{14}\end{aligned}\tag{5.100}$$

where β_0 represents the common propagation constant. The four super-modes defined in Eq. (5.100) can be sequentially denoted as super-mode 1 through super-mode 4, and the inter super-mode coupling coefficients are symbolized as K_{21} K_{31} , K_{41} ... etc.

Table 5.3 The detailed parameters of the 4-core fiber in [72]

Parameters	Value
Number of cores	4
Core radius	2.5 μm
Cladding diameter	125.7 μm
Core pitch 1	5 μm
Core pitch 2	8.66 μm
Cladding index	1.45
Core index difference	0.31%

This table is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

As stated earlier, polarization is considered in the Monte Carlo simulations with the vector model [71]:

$$\frac{d\vec{E}_m}{dz} = i\beta_m\vec{E}_m + \sum_{n=1}^{N-1} i\beta_{mn}\vec{E}_n + \sum_{n=1}^N i\kappa_{mn}f_{mn}(z)\vec{E}_n + \frac{i\vec{\alpha}_m}{2} \cdot \vec{\sigma} \vec{E}_m, \quad (5.101)$$

where $\vec{E}_m$ and $\vec{\alpha}_m$ are the Jones vector and the random birefringence vector of the m th core, $\vec{\sigma}$ is the Pauli matrices vector. The correlation function of $\vec{\alpha}_m$ fulfills [71]

$$\langle \vec{\alpha}_n(z+z') \cdot \vec{\alpha}_n(z') \rangle = \left(\frac{2\pi}{L_B} \right)^2 \exp\left(-\frac{|z|}{L_C} \right), \quad (5.102)$$

where L_C is the birefringence correlation length, a parameter that quantifies the rate of attenuation in the birefringence autocorrelation function, and L_B denotes the birefringence beat length, which characterizes the strength of the birefringence effect.

It is presumed that L_C is 1 m and L_B is 10 m, with these parameters being employed and analyzed in the previously published sections [71]. The random fluctuations in the coupling coefficients are considered to exhibit a root mean square (rms) value of $0.002\beta_{12}$ for the two-core fiber, and correspondingly, $0.002\beta_{12}$, $0.002\beta_{13}$, and $0.002\beta_{14}$ for the four-core fiber. To calculate the average values of power and power correlation terms, Monte Carlo simulations are conducted with 1000 iterations.

Given that the birefringence correlation length is significantly shorter compared to the beating length for deterministic coupling of super-modes (on the order of millimeters, as mentioned in [72]), our proposed scalar analytical model, which disregards the polarization's influence on crosstalk, specifically Eqs. (5.92–5.93b), can be applied effectively.

Initially, the two-core fiber is simulated. As discussed in Sect. 5.2, the random coupling correlation length l_c is intimately linked to the crosstalk between super-modes, and it is assumed to be 1 m. Assuming that unit power is injected into the symmetrical super-mode, its crosstalk to the anti-symmetrical super-mode is computed and presented in Fig. 5.16a, b. The dash-dotted line, dashed line, and solid line represent the outcomes from Monte Carlo simulations, the simplified model from Eqs. (5.90, 5.95, and 5.97), and the analytical model, respectively. The figure clearly demonstrates that both the analytical model and the simplified formula accurately predict the average crosstalk power and its standard deviation, which closely matches the Monte Carlo simulations. Furthermore, it reveals that the crosstalk level can fluctuate by up to 3 dB, which is the crucial information for system design and analysis, particularly when evaluating the receiver's signal-to-noise ratio (SNR) budget.

Subsequently, we investigate the influence of l_c on crosstalk by varying its value to 1, 4, 7, and 10 m, respectively, with a fixed fiber length of 100 m. The graphs in Fig. 5.17a, b indicate that the analytical model accurately predicts both the average crosstalk power and its standard deviation compared to Monte Carlo simulations, showcasing the model's robustness. Moreover, the figures suggest that both the average crosstalk power and its standard deviation decrease with an increase in l_c ,

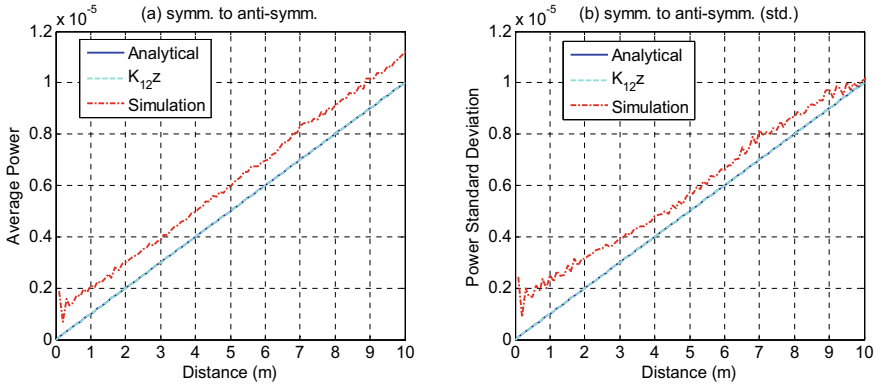


Fig. 5.16 The evolution of the crosstalk power and power standard deviation from the symmetrical super-mode to the anti-symmetrical super-mode. The random coupling correlation length $l_c = 1$ m. The unit power is injected on the symmetrical super-mode. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.90, 5.95, and 5.97), and the analytical model. This figure is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

consistent with Eq. (5.99). The deviation between the curves arises from the underlying assumption in the theory: the random coupling correlation length is significantly shorter than the propagation distance, enabling the random process to be modeled as the Brownian motion. However, when the propagation distance is short or comparable to the random coupling correlation length, discrepancy from the theory emerge, notably at the initial stages of propagation. Figure 5.17c, d quantitatively analyze these discrepancies, revealing that they increase as l_c increases. Conversely, as the propagation distance extends, the discrepancies diminish.

Afterwards, we simulate a four-core fiber, initially assuming a random coupling correlation length l_c of 1 m. Assuming unit power injection into the LP₂₁-like super-mode, we evaluate the crosstalk to the other super-modes. The results, presented in Fig. 5.18a–c, demonstrate that the proposed analytical model and simplified formulas, specifically Eqs. (5.90 and 5.95), provide highly satisfactory predictions when compared to Monte Carlo simulations. Notably, the crosstalk between the LP₂₁-like mode and the LP_{11b}-like mode is significantly higher than that with the other two super-modes. This phenomenon arises from the smaller difference in propagation constants between the LP₂₁-like and LP_{11b}-like modes. Figure 5.18d illustrates the evolution of the standard deviation for the crosstalk between these two modes, showing that the proposed analytical model and simplified formulas, including Eq. (5.97), exhibit exceptional accuracy in assessing the power variation of super-mode crosstalk.

It is noteworthy that the simplified formulas, specifically Eqs. (5.90, 5.95, and 5.97), exhibit deviations from the analytical results under conditions of high crosstalk levels, such as -10 dB. Consequently, the analytical solution presents a more precise method for predicting the ISMXT power and its standard deviation.

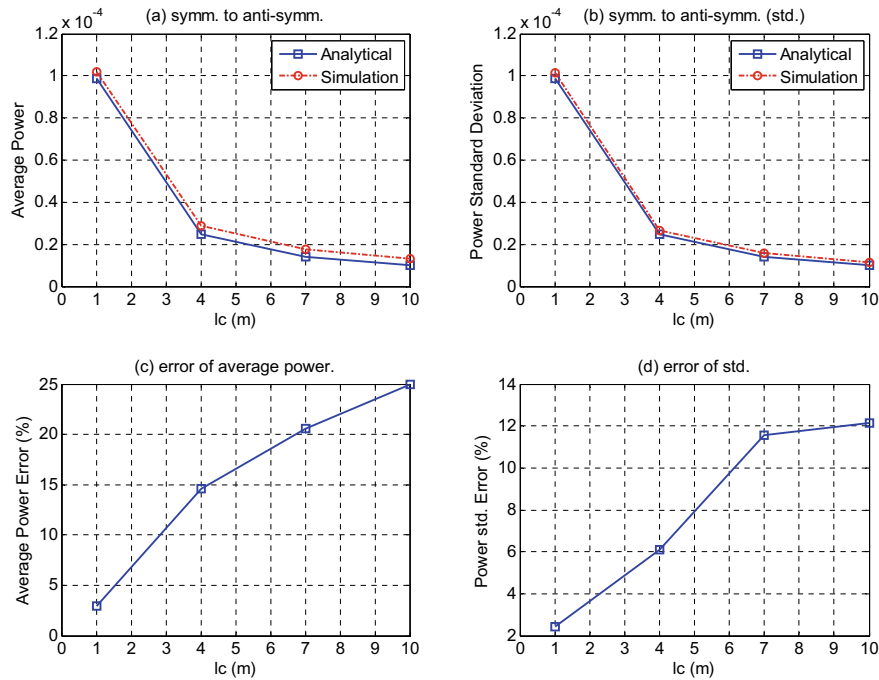


Fig. 5.17 The change of the crosstalk power and the power standard deviation from the symmetrical super-mode to the anti-symmetrical super-mode with respect to the value of l_c . The unit power is injected on the symmetrical super-mode and the fiber length is 100 m. The results are computed by the Monte Carlo simulations and the analytical model. The discrepancy between the two models is shown in **c** and **d**. This figure is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

Thereafter, we assume the correlation length l_c as 10 m, with the continued assumption of unit power injection into the LP21-like super-mode. The results of crosstalk are presented in Fig. 5.19a–c. The match between Monte Carlo simulations, the simplified formulas (Eqs. 5.90 and 5.95), and the analytical model further validates the model’s efficacy. When compared to the results in Fig. 5.18a–c, it becomes evident that the crosstalk level, which is inversely proportional to the random correlation length l_c , has decreased by a factor of 10. This implies that manufacturing MCFs with high consistency (i.e., longer correlation lengths for variations) can effectively mitigate crosstalk between super-modes. Notably, the final crosstalk level between the LP21-like mode and the LP11b-like mode in Fig. 5.19a is approximately -23 dB/km, aligning with measurement results reported in [72], suggesting that the parameter configuration might closely resemble the actual parameters employed during fabrication [72].

Figure 5.19d illustrates the evolution of the standard deviation for the crosstalk between the LP21-like mode and the LP11b-like mode, given a correlation length l_c of 10 m. The close match among the three lines reinforces the validity of the proposed

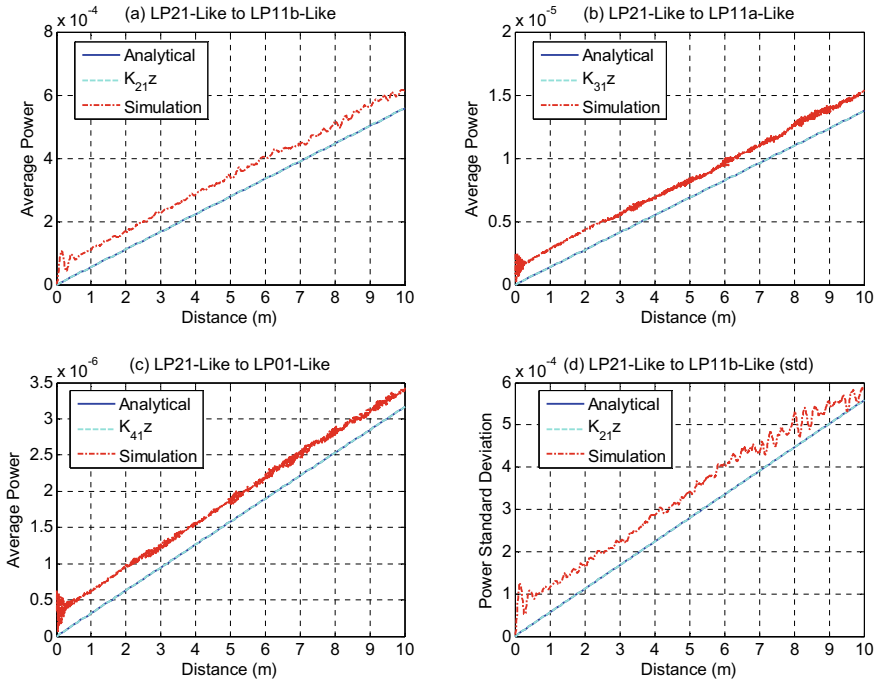


Fig. 5.18 The evolution of the crosstalk power from the LP21-like mode to the **a** LP11b-like mode, **b** LP11a-like mode, and **c** LP01-like mode. **d** The evolution of the crosstalk power standard deviation from the LP21-like mode to the LP11b-like mode. The random coupling correlation length $l_c = 1$ m. The unit power is injected on LP21-like mode. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.90, 5.95, and 5.97), and the analytical model. This figure is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

model. As the crosstalk level diminishes by a factor of 10 with the increase in l_c , the standard deviation of the crosstalk also experiences a corresponding decrease.

It is important to highlight that the presented theoretical framework is versatile and can be applied to a diverse range of multi-core fibers (MCFs) designed for super-mode transmission, regardless of their core arrangement. The specific case of a rectangular 4-core fiber, as previously demonstrated, exemplifies its use in achieving MIMO-less transmission by leveraging non-degenerate super-modes. Essentially, as long as the super-modes remain non-degenerate, the theory provides accurate predictions for the crosstalk between them. In the context of standard non-rectangular MCFs, where cores may be arranged in a square configuration, degenerate super-modes may arise. While the theory remains capable of predicting crosstalk between super-modes with distinct propagation constants, it falls short in predicting crosstalk between those sharing identical propagation constants. Fortunately, crosstalk between super-modes featuring different propagation constants is typically of greater significance and concern [73].

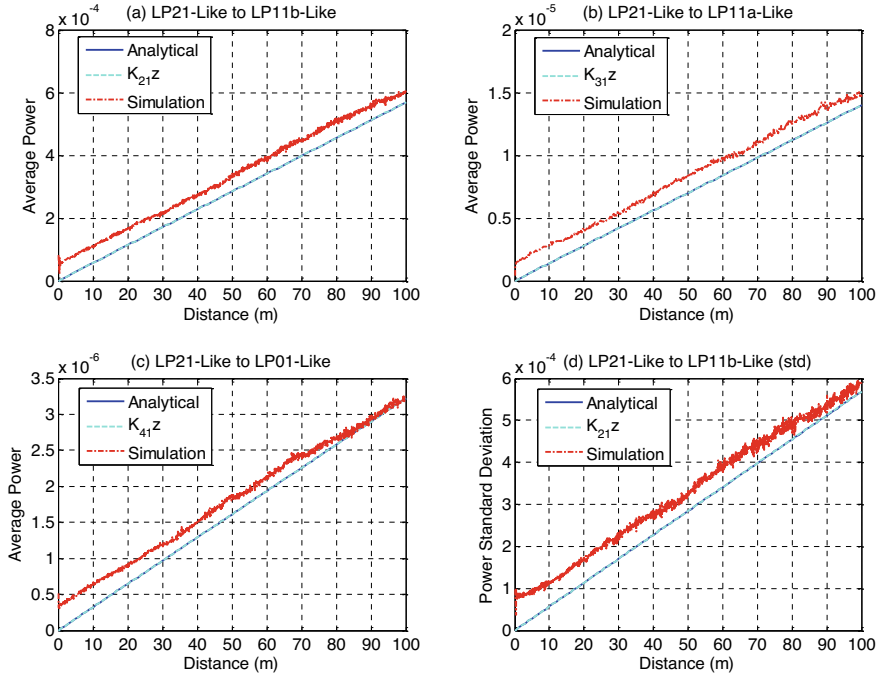


Fig. 5.19 The evolution of the crosstalk power from the LP21-like mode to the **a** LP11b-like mode, **b** LP11a-like mode, and **c** LP01-like mode. **d** The evolution of the crosstalk power standard deviation from the LP21-like mode to the LP11b-like mode. The random coupling correlation length $l_c = 10$ m. The unit power is injected on LP21-like mode. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.90, 5.95, and 5.97), and the analytical model. This figure is reprinted/adapted with permission from [Opt. Express 30, 4833–4844 (2022)] © The Optical Society of America [55]

5.3.4 Summary

An analytical methodology for assessing super-mode crosstalk in tightly confined MCFs is introduced, and its precision is validated through Monte Carlo simulations. The findings indicate that crosstalk can be mitigated by enhancing the difference in propagation constants and lengthening the correlation length of random variations. When deterministic coupling strength is pronounced, it leads to a wider gap in propagation constants, consequently anticipating a notably shorter coupling beat length. This implies minimal alteration in signal polarization during the coupling process, suppressing the influence of PMD. Furthermore, manufacturing MCFs with high consistency to get long correlation lengths for random variations will also lead to reduced crosstalk levels. The theoretical insights presented in this study offer valuable guidance for the design and analysis of MCFs.

5.4 On the Variance of Crosstalk in Multi-core Fibers: With and Without the Polarization Effect

In this section, the analysis on the multicore fiber based on the random coupled theory continues. The variance of the inter core crosstalk (ICXT) in multi-core fibers (MCFs) is studied with explicit formulas, which suggest the crosstalk power variance equals the square or the half square of the crosstalk power when the PMD effect is considered or not considered. The analysis is based on the stochastic differential equations (SDEs) derived from the random coupled mode theory (CMT). Monte Carlo simulations are performed and the accuracy of the predictions is verified. The Summary presented here is beneficial for the design and analysis of the general MCFs and the polarization maintaining MCFs (PMMCFs).

5.4.1 Introduction

Recently, multicore fibers (MCFs) have gained significant attentions in the fiber industry due to their capability to further increase the transmission capacity without increasing the size of fiber cross section as compared with that of the single mode fibers [25, 56–58].

Crosstalk is the one of the important issues for MCF design and analysis. The inter core crosstalk (ICXT) is not only dependent on the coupling strength between the cores, but also impacted by the random variations caused by either the random phase change or the random polarization rotation in the cores [61–71, 78]. Numerous studies on the topic have been presented. The theoretical contribution started from Marcuse [13], who proposed the well-known coupled power theory, which turned the problem of random coupling into a set of ordinary differential equations (ODEs). The subsequent studies are mostly based on the theory, because it can provide a simple explicit formula for the crosstalk characterization [61–71, 78]. While most of the studies ignore the polarization mode dispersion (PMD) induced random coupling, Antonelli et al. [71] proposed a full analytical theory which incorporated the PMD effect into the crosstalk evaluation formula.

While the existing theory accurately deals with the power evolution in the MCFs, the power variance lacks a systematic treatment. The variance of the ICXT power inside the cores is essential for the receiver signal to noise ratio (SNR) budget analysis, which is crucial from the perspective of network architecture design. Monte Carlo simulations are the main approaches to obtain the crosstalk power variance, which is, however, extremely time consuming while dealing with the MCFs with a short random variation correlation length. One may also assume the ICXT to be the random chi square distribution, and calculate the variance with respect to the average power [71]. However, it is lack of physical implications and the accuracy of the model is highly dependent on the experience of the researcher to choose the degree of freedom

for chi square distribution. Furthermore, it is not clear how the polarization effect impacts the ICXT variance of the MCFs.

In our previous work, the ICXT variance was studied [77]. However, no explicit formulas are derived in [77]. In [55], the crosstalk power and the variance for the super-modes in MCFs are given; however, the polarization effects are not included because the super-modes in the MCFs have large index differences and this suppresses the impact of PMD. The suppression of the PMD impact is caused by the fact that the signal polarization state barely changes within the coupling beat length. Therefore, the ICXT variance for the MCFs with and without the polarization coupling were not clearly identified and analyzed. In this work, based on the vector stochastic coupled mode theory, we derive the coupled power equation and the coupled 2nd order power moment equation with and without the PMD effect. Explicit formulas are given, which suggest the crosstalk power variance equals the square or the half square of the crosstalk power when the PMD effect is considered or not considered. Both Summaries are useful in the real applications. For the former, it is useful for the ICXT analysis in the general MCFs. The latter prediction will be accurate for the polarization maintaining MCFs (PMMCFs) [79–81], which are designed to enable non-DSP-aided MIMO transmission. Monte Carlosimulations are performed to verify the proposed assertions.

5.4.2 Theory

If the PMD is included in the random couple mode theory, the following equation can be used to characterize the vector beam propagation in the MCFs [71]:

$$\frac{d\vec{E}_m}{dz} = i\beta_m\vec{E}_m + \sum_{n=1}^N i\kappa_{mn}f_{mn}(z)\vec{E}_n + \frac{i\vec{\alpha}_m}{2} \cdot \vec{\sigma}\vec{E}_m, \quad (5.103)$$

where $\vec{E}_m$, $\vec{\alpha}_m$ and β_m are the Jones vector, the random birefringence vector and the propagation constant of the mode in the m th core, and $\vec{\sigma}$ is a matrix vector whose elements are the Pauli matrices. κ_{mn} and f_{mn} stand for the coupling amplitude and the z directional random variation process between the m th and the n th cores, where f_{mn} is a filtered color noise with the correlation function as [61, 62, 68, 71]

$$\begin{aligned} \langle f_{mn}(z+z')f_{mn}^*(z') \rangle &= \eta_{mn}(z), \\ \eta_{mn}(z) &= \eta(z). \end{aligned} \quad (5.104)$$

Without loss of generality, it is assumed in Eq. (5.104) that f_{mn} are identically distributed independent (i.i.d) variables. Taking the following transformation [71, 77],

$$\begin{aligned}\mathbf{U}_m \vec{A}_m &= \exp(-i\beta_m z) \vec{E}_m, \\ \frac{d\mathbf{U}_m}{dz} &= \frac{i\vec{\alpha}_m}{2} \cdot \vec{\sigma} \mathbf{U}_m,\end{aligned}\tag{5.105}$$

one may convert Eq. (5.103) into the following [77]:

$$\begin{aligned}\frac{d\vec{A}_m}{dz} &= \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n \vec{A}_n, \\ F_{mn}(z) &= \kappa_{mn} f_{mn}(z) \exp(-j(\beta_m - \beta_n)z),\end{aligned}\tag{5.106}$$

where the matrix $\mathbf{U}_m$ has the following property [71, 77]

$$\begin{aligned}\mathbf{U}_n(0) &= \mathbf{I}, \\ \mathbf{U}_n(z) &= \mathbf{U}_n(z_1) \mathbf{U}_n(z - z_1), \\ \mathbf{U}_n^{-1}(z) &= \mathbf{U}_n^H(z) = \mathbf{U}_n(-z), \\ \langle \mathbf{U}_n(z) \mathbf{U}_n^H(z') \rangle &= \langle \mathbf{U}_n(z - z') \rangle = \xi_n(z - z') \mathbf{I}.\end{aligned}\tag{5.107}$$

ξ_n stands for the correlation function for matrix $\mathbf{U}_n$. In case the polarization effect can be ignored (e.g., in the PMMCFs), $\xi_m = \xi_n = 1$.

Defining [71, 77],

$$K_{mn} = |\kappa_{mn}|^2 \mathbf{F} \mathbf{T} (\eta \xi_m \xi_n) (\beta_m - \beta_n),\tag{5.108}$$

we may have the following property for the coupling matrix in Eq. (5.106):

$$\begin{aligned}\left\langle F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n)^H dz \right\rangle &= \mathbf{I} K_{mn} dz, \\ \left\langle F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n) dz \right\rangle &= \mathbf{0}, \\ \left\langle F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz (F_{m'n'} \mathbf{U}_{m'}^{-1} \mathbf{U}_{n'})^H dz \right\rangle &= \mathbf{0} (m \neq m', n \neq n'), \\ \left\langle F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz (F_{m'n'} \mathbf{U}_{m'}^{-1} \mathbf{U}_{n'}) dz \right\rangle &= \mathbf{0} (m \neq m', n \neq n').\end{aligned}\tag{5.109}$$

The brief proof is as follows. Based on Eqs. (5.106 and 5.107), we have

$$\begin{aligned}
& \left\langle F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n)^H dz \right\rangle \\
&= \left\langle \int_z^{z+dz} F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz' \int_z^{z+dz} F_{mn}^* \mathbf{U}_n^H (\mathbf{U}_m^{-1})^H dz'' \right\rangle \\
&= |\kappa_{mn}|^2 \mathbf{I} \int_z^{z+dz} dz' \int_z^{z+dz} \eta_{mn}(z' - z'') \xi_m(z'' - z') \xi_n(z' - z'') \exp(-j(\beta_m - \beta_n)(z' - z'')) dz'' \\
&\approx |\kappa_{mn}|^2 \mathbf{F} \mathbf{T} (\eta_{mn} \xi_m \xi_n) (\beta_m - \beta_n) dz \mathbf{I},
\end{aligned} \tag{5.110}$$

which is the first sub-equation in Eq. (5.109). For the other sub-equations in Eq. (5.109), they all have the fast-oscillating term $\exp(j(\beta_m - \beta_n)z'')$ in the integral, which tends to be 0 during the integration. Therefore, we have Eq. (5.109).

Equation (5.106) is the Stratonovich sense stochastic differential equation (SDE), and it should be converted to the Ito sense SDE. Using Eq. (5.109), such conversion can be done:

$$\begin{aligned}
d\vec{A}_m &= \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \left(\vec{A}_n(z) + \frac{1}{2} d\vec{A}_n \right) \\
&= \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n + \frac{1}{2} \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz d\vec{A}_n \\
&= \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n - \frac{1}{2} \sum_{n=1}^N F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \sum_{m'=1}^N F_{nm'} \mathbf{U}_n^{-1} \mathbf{U}_{m'} dz d\vec{A}_{m'} \\
&= \sum_{n=1}^N iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n - \frac{1}{2} \sum_{n=1}^N K_{mn} \vec{A}_m dz,
\end{aligned} \tag{5.111}$$

Defining the power of the m th core as $P_m = \vec{A}_m^* \cdot \vec{A}_m$, and we have the following SDE for the power evolution of the m th core:

$$\begin{aligned}
dP_m &= d(\vec{A}_m^* \cdot \vec{A}_m) \\
&= d\vec{A}_m^* \cdot \vec{A}_m + \vec{A}_m^* \cdot d\vec{A}_m + d\vec{A}_m^* \cdot d\vec{A}_m \\
&= \sum_{n=1}^N K_{mn} (P_n - P_m) dz + \sum_{n=1}^N \left(iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n \right)^* \cdot \vec{A}_m \\
&\quad + \sum_{n=1}^N \vec{A}_m^* \cdot \left(iF_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n \right).
\end{aligned} \tag{5.112}$$

Taking the average on Eq. (5.112), one gets the well-known ODE, i.e., the coupled power equation [13, 77]

$$\frac{d\langle P_m \rangle}{dz} = \sum_{n=1}^N K_{mn} (\langle P_n \rangle - \langle P_m \rangle) \quad (5.113)$$

Based on Eq. (5.112), one has:

$$\begin{aligned} d\langle P_m^2 \rangle &= \langle 2P_m dP_m \rangle + \langle dP_m dP_m \rangle \\ &= 2 \sum_{n=1}^N K_{mn} P_m (P_n - P_m) dz \\ &\quad + 2 \left\langle \sum_{n=1}^N \left((F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)^* \cdot \vec{A}_m \right) (\vec{A}_m^* \cdot (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)) \right\rangle \\ &\quad - 2 \left\langle \left((F_{mm} \mathbf{U}_m^{-1} \mathbf{U}_m dz \vec{A}_m)^* \cdot \vec{A}_m \right)^2 \right\rangle \end{aligned} \quad (5.114)$$

In the case the Jones vectors $\vec{A}_n$ and $\vec{A}_m$ are fully de-correlated (which is usually the case when PMD is taken into account), we have

$$\begin{aligned} \left\langle \sum_{n=1}^N \left((F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)^* \cdot \vec{A}_m \right) (\vec{A}_m^* \cdot (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)) \right\rangle &\approx \frac{1}{2} \sum_{n=1}^N K_{mn} P_m P_n dz, \\ \left\langle \left((F_{mm} \mathbf{U}_m^{-1} \mathbf{U}_m dz \vec{A}_m)^* \cdot \vec{A}_m \right)^2 \right\rangle &\approx \frac{1}{2} K_{mm} P_m^2 dz. \end{aligned} \quad (5.115)$$

The coupled ODEs for the powers and the power 2nd order moments in the cores can be derived as:

$$\begin{aligned} \frac{d\langle P_m^2 \rangle}{dz} &= -2 \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_m^2 \rangle + 3 \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_m P_{m'} \rangle, \\ \frac{d\langle P_m P_n \rangle}{dz} &= \left(- \sum_{m'=1, m' \neq m}^N K_{mm'} - \sum_{n'=1, n' \neq n}^N K_{nn'} \right) \langle P_m P_n \rangle + \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_n P_{m'} \rangle \\ &\quad + \sum_{n'=1, n' \neq n}^N K_{nn'} \langle P_m P_{n'} \rangle - K_{mn} \langle P_m P_n \rangle. \end{aligned} \quad (5.116)$$

In the case the Jones vectors $\vec{A}_n$ and $\vec{A}_m$ are co-polarized and the polarization rotation matrices are the identity matrix (which means the polarization effect is ignored, e.g., in the PMMCFs), we have

$$\begin{aligned}
\left\langle \sum_{n=1}^N \left((F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)^* \cdot \vec{A}_m \right) (\vec{A}_m^* \cdot (F_{mn} \mathbf{U}_m^{-1} \mathbf{U}_n dz \vec{A}_n)) \right\rangle &= \sum_{n=1}^N K_{mn} P_m P_n dz, \\
\left\langle \left((F_{mm} \mathbf{U}_m^{-1} \mathbf{U}_m dz \vec{A}_m)^* \cdot \vec{A}_m \right)^2 \right\rangle &= K_{mm} P_m^2 dz.
\end{aligned} \tag{5.117}$$

The coupled ODEs for the powers and power 2nd order moments should be modified as

$$\begin{aligned}
\frac{d\langle P_m^2 \rangle}{dz} &= -2 \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_m^2 \rangle + 4 \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_m P_{m'} \rangle, \\
\frac{d\langle P_m P_n \rangle}{dz} &= \left(- \sum_{m'=1, m' \neq m}^N K_{mm'} - \sum_{n'=1, n' \neq n}^N K_{nn'} \right) \langle P_m P_n \rangle + \sum_{m'=1, m' \neq m}^N K_{mm'} \langle P_n P_{m'} \rangle \\
&\quad + \sum_{n'=1, n' \neq n}^N K_{nn'} \langle P_m P_{n'} \rangle - 2K_{mn} \langle P_m P_n \rangle.
\end{aligned} \tag{5.118}$$

Equation (5.112) together with Eq. (5.116) or Eq. (5.118) can be used for the crosstalk and its variance evaluation with and without the polarization effect. In the case of weak crosstalk, explicit formulas for the crosstalk and the crosstalk variance can be derived. Considering the ICXT from core n to core m , we neglect the power from other cores in Eq. (5.112) and compute it as:

$$\begin{aligned}
\langle P_m(z) \rangle &\approx \langle P_n(0) \rangle K_{mn} z, \\
\langle ICXT_{mn} \rangle &= \left\langle \frac{P_m(z)}{P_n(0)} \right\rangle \approx K_{mn} z,
\end{aligned} \tag{5.119}$$

which is the common-sense crosstalk formula for the MCFs [13, 55, 61–71, 78]. In the same way, we may compute the 2nd order power moments from Eq. (5.116) with the fully de-correlated Jones vectors as

$$\begin{cases} \langle P_m P_n \rangle \approx K_{mn} \int_0^z \langle P_n^2 \rangle dz \approx K_{mn} z \langle P_n(0)^2 \rangle, \\ \langle P_m^2 \rangle \approx 3K_{mn} \int_0^z \langle P_m P_n \rangle dz = \frac{3}{2} (K_{mn} z)^2 \langle P_n(0)^2 \rangle. \end{cases} \tag{5.120}$$

Considering the fact that

$$\langle P_n(0) \rangle^2 = \langle P_n(0)^2 \rangle, \tag{5.121}$$

we have the variance of $ICXT_{mn}$ in the presence of PMD as

$$var(ICXT_{mn}) = var\left(\frac{P_m^2}{\langle P_n(0)^2 \rangle}\right) \approx \frac{1}{2}(K_{mn}z)^2, \quad (5.122)$$

and the standard deviation can be calculated as

$$std(ICXT_{mn}) = \frac{1}{\sqrt{2}}K_{mn}z. \quad (5.123)$$

Similarly, the variance and the standard deviation of the ICXT without the impact of PMD can be calculated from Eq. (5.118), which is similar to the formulas for the crosstalk variance and the standard deviation for the super-modes within the MCFs without considering the polarization effect [55]

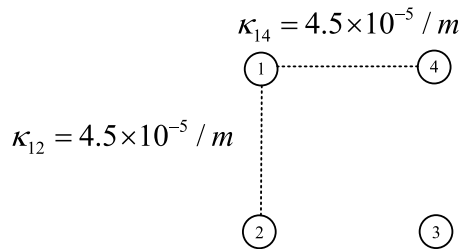
$$\begin{aligned} var(ICXT_{mn}) &\approx (K_{mn}z)^2, \\ std(ICXT_{mn}) &\approx K_{mn}z. \end{aligned} \quad (5.124)$$

5.4.3 Results and Discussions

Monte Carlo simulations have been conducted to verify the proposed analytical model and the explicit formulas (Eqs. 5.119, 5.123, and 5.124) provided above. A four-core MCF is simulated with its cross section shown in Fig. 5.20. The fiber has the random coupling strength as $4.5 \times 10^{-5}/m$ between the neighboring cores and the coupling between the diagonal cores is neglected, which is the same as the case in [71].

In the first simulation setup, the four core MCF is assumed to be affected by the random PMD effect, with the random birefringence vector of the m th core $\vec{\alpha}_m$ fulfilling

Fig. 5.20 The four-core fiber cross section



$$\langle \vec{\alpha}_n(z + z') \cdot \vec{\alpha}_n(z') \rangle = \left(\frac{2\pi}{L_B} \right)^2 \exp\left(-\frac{|z|}{L_C}\right), \quad (5.125)$$

where L_C is the polarization correlation length and L_B is the polarization beat length, and their values are assumed as 1 and 10 m respectively. The random coupling correlation length l_c is assumed as 10 m. 1000 realizations are evaluated in the Monte Carlo simulations to compute the average power and the power standard deviation for the ICXT.

The computed results are shown in Fig. 5.21, which includes the results by the Monte Carlo simulations, the explicit formulas (Eqs. 5.119 and 5.123) and the analytical solutions of Eqs. (5.113 and 5.116). The agreement between the curves indicates the validity of the variance model with the polarization effect. It is clearly shown in the figure that the power standard deviation of the ICXT is indeed the average power multiplied by the factor of $\sqrt{1/2}$, which is caused by the depolarization effect for the Jones vectors induced by the random birefringence.

The proposed explicit formula is equally accurate when the correlation length parameter changes. In Fig. 5.22, the random coupling correlation length l_c is assumed to be 0.5 m. The well-match between the curves suggests the model is robust with the change of the parameters. The scaling factor of $\sqrt{1/2}$ still exists between the average power and the power standard deviation.

The case without the polarization effect is studied afterwards. Without loss of generality, the same MCF presented in Fig. 5.20 is considered as a PMMCF, which can neglect the polarization effect during the ICXT analysis. The random correlation length l_c is assumed to be 10 and 0.5 m respectively. The results by the Monte Carlo simulations, the explicit formulas (Eqs. 5.119 and 5.124), and the analytical solutions of Eqs. (5.113 and 5.118) are shown in Figs. 5.23 and 5.24. The curves

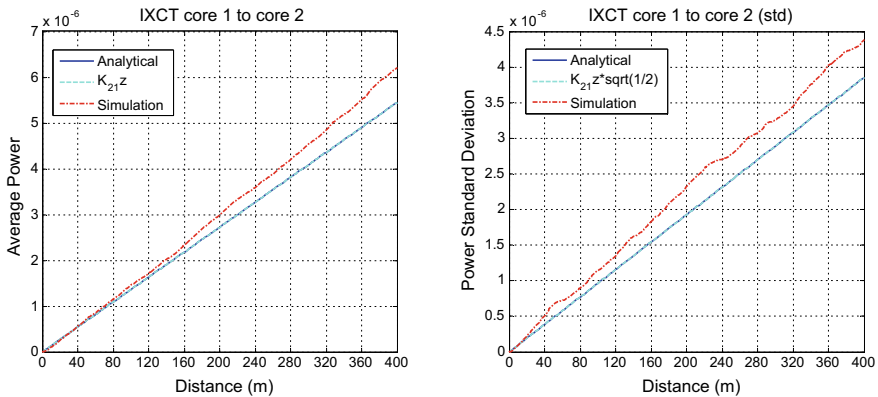


Fig. 5.21 The evolution of the crosstalk average power and the power standard deviation in the presence of the polarization effect. The random coupling correlation length $l_c = 10$ m. The unit power is injected into core 1. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.119 and 5.123), and the analytical model

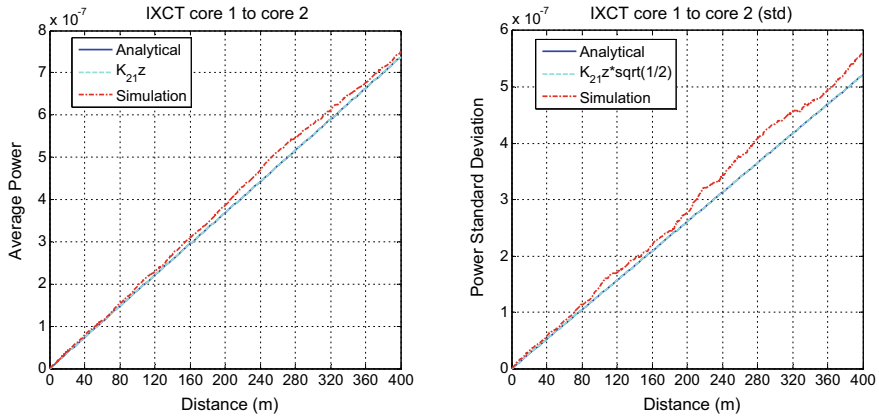


Fig. 5.22 The evolution of the crosstalk average power and the power standard deviation in the presence of the polarization effect. The random coupling correlation length $l_c = 0.5$ m. The unit power is injected into core 1. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.119 and 5.123), and the analytical model

clearly demonstrate that the scaling factor between the average power and the power standard deviation for ICXT has changed from $\sqrt{1/2}$ to 1. These results confirm the predictions in Sect. 5.2.

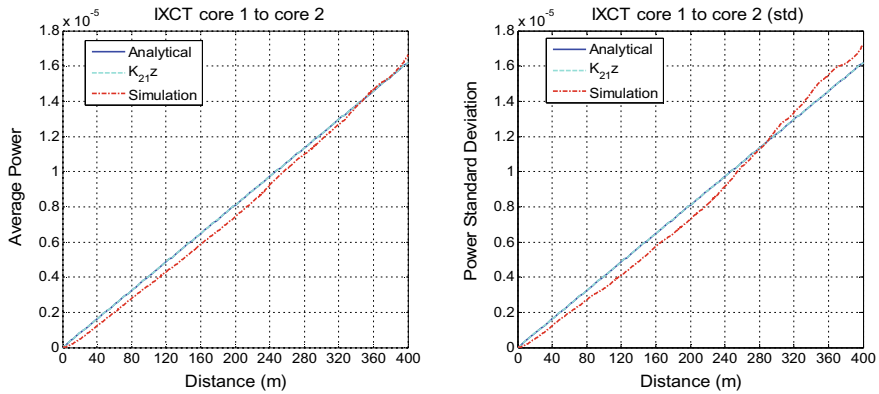


Fig. 5.23 The evolution of the crosstalk average power and the power standard deviation without the polarization effect. The random coupling correlation length $l_c = 10$ m. The unit power is injected into core 1. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.119 and 5.124), and the analytical model

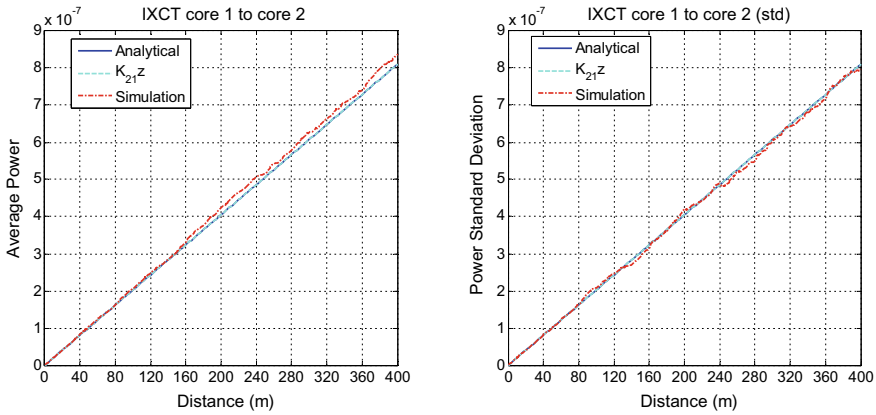


Fig. 5.24 The evolution of the crosstalk average power and the power standard deviation without the polarization effect. The random coupling correlation length $l_c = 0.5$ m. The unit power is injected into core 1. The results are computed by the Monte Carlo simulations, the explicit formulas in Eqs. (5.119 and 5.124), and the analytical model

5.4.4 Summary

The variance for the crosstalk in MCFs is discussed with and without the polarization effect. The analysis is based on the SDEs derived from the vector coupled mode equations for the MCFs. ODEs for the powers and the 2nd order power moments are derived. Explicit formulas are provided, which suggests crosstalk power variance equals the square or the half square of the crosstalk power when the PMD effect is considered or not considered. Monte Carlo simulations are performed, which verifies the accuracy of the above predictions. The Summary in this work can be used for quantitative evaluation of the crosstalk variance during the MCF design and the system analysis.

5.5 A Coupled Mode Theory for Orbital Angular Momentum Modes Transmission in the Presence of Atmosphere Turbulence

Part of the section reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82].

As is shown in the previous section, the random coupled mode theory achieves accurate statistical analysis for the modal propagation in the multimode and multicore waveguides. The theory can be readily extended for the free space wave propagation. In this section, we present that the random coupled mode theory can be used

to study the orbital angular momentum (OAM) modes transmission in the presence of atmosphere turbulence. The Laguerre–Gauss (LG) beams are assumed to propagate in free space and inevitably encounter interactions mediated by the atmosphere’s random refractive index fluctuations. The intricate interplay between these LG beams is quantitatively captured by a system of coupled average power equations like Marcuse’s Coupled Power Equation (CPE) initially devised for optical waveguide analysis. Within this framework, both the coupling coefficients and modal radiation losses can be analytically assessed. Notably, both the precise solution and its first-order approximation closely align with established data and Monte Carlo simulation outcomes, exhibiting high accuracy. As such, the CPE and its approximate analytical formulations serve as potent instruments for dissecting the evolution of OAM beams amidst atmospheric turbulence.

5.5.1 Introduction

Optical Orbital Angular Momentum (OAM) multiplexing stands as a highly promising approach for achieving space division multiplexing (SDM) in optical communication systems [83–86]. This technique implements OAM modes with distinct topological charges, as the carriers for information transmission, thereby significantly enhancing the overall system capacity.

In an ideally radial symmetric environment, the topological charge of an OAM mode remains intact. The inherent orthogonality between OAM modes facilitates straightforward multiplexing and de-multiplexing among channels. Given that most current experiments of free-space OAM transmission have been conducted over short distances [83–85], this orthogonality is impeccably preserved, rendering the crosstalk between the OAM modes during transmission negligible. Consequently, the primary limitation comes from the OAM multiplexing and de-multiplexing components.

Nevertheless, as the transmission distance prolongs, the variations in refractive index due to uneven distributions of temperature and pressure within the atmosphere cannot be overlooked and become significant factors. It is well-established that atmospheric turbulence can introduce crosstalk during the propagation of OAM modes, prompting numerous studies aimed at quantifying its impact. [87–91]. Paterson [87] and Tyler [89] provided analytical formulas for the weak turbulence case, which assumed the impact of the index variation can be modelled by the phase variations of the beam. Rodenburg [90] experimentally validated the statistical theory of phase perturbations. Zhang et al. presented comparable findings for Whittaker-Gaussian (WG) beams transmitting under turbulent atmospheric conditions [91]. However, as highlighted in these pioneering works [90], the phase perturbation model may lose validity in long-distance point-to-point communications due to the complexity of random scattering in such environments, rendering the model inaccurate [90]. To study the phenomenon in a more general sense, To address this, J. A. Anguita employed the split-step Fourier method (SSFM) to simulate the propagation process, providing accurate characterizations for long-distance beam propagation in media

with significant refractive index variations [88]. Nevertheless, this method is computationally intensive as the inherent randomness of atmospheric turbulence necessitates Monte Carlo simulations, involving numerous iterations to determine average values. Despite the extensive analytical and numerical research conducted on this subject [87–91], there remains a pressing need for an efficient analytical tool capable of studying strong atmospheric turbulence during relatively long-distance free-space propagation.

In this section, we introduce a set of coupled power equations (CPE) tailored for OAM mode transmission through atmospheric turbulence, which resembles Marcuse’s coupled power theory [5–7] for optical waveguides with index perturbations. We demonstrate that the coupling coefficients and mode radiation losses within this CPE framework can be analytically derived. The CPE comprises a system of ordinary differential equations (ODEs) that can be efficiently solved using numerical integration methods, such as the Runge–Kutta method. By performing the numerical integration just once, we significantly reduce computational costs compared to Monte Carlo simulations. Furthermore, given that LG modes undergo relatively small variations within the Rayleigh range, the coupling coefficients and radiation losses can be considered constant. Consequently, the first-order approximation to the CPE, which assumes these coefficients and losses remain constant, provides a reasonable estimate. As this first-order approximation is analytically solvable, it enhances the CPE’s utility as a powerful tool for analyzing long-distance OAM mode transmission in the turbulent atmospheric conditions. The validity of our proposed CPE is corroborated through comparisons with published data in [88] and Monte Carlo simulation outcomes, underscoring its effectiveness and accuracy.

5.5.2 The Coupled Power Equations

In order to find the CPE analogous to the Marcuse’ equations [5–7], the coupled mode equations for the OAM modes are formulated under the condition of index perturbation.

The scalar wave equation is adopted in this case [87–92]

$$\nabla^2 \varphi + k^2 \varphi = 0 \quad (5.126)$$

where k is the wave number under the atmosphere turbulence, which is formulated as [92]

$$(1 + \delta n)k_0 \quad (5.127)$$

where k_0 denotes the free space wave number, and δn denotes the refractive index turbulence. The variation in the refractive index can be adjusted to behave as a zero-mean random process across the three-dimensional space by subtracting the average value of δn from the overall refractive index fluctuations, i.e.

$$\Delta n = \delta n - \langle \delta n \rangle \quad (5.128)$$

Hence, we have

$$\begin{aligned} k &= (n + \Delta n)k_0 \\ n &= 1 + \langle \delta n \rangle \end{aligned} \quad (5.129)$$

Because of the small index variation (in the order of 10^{-6}) [92], the average term $\langle \delta n \rangle$ is negligible. Disregarding the higher order terms in the equation, one has [92]

$$\nabla^2 \varphi + k_0^2 (1 + 2\Delta n) \varphi = 0 \quad (5.130)$$

Without the influence of atmospheric turbulence, Laguerre–Gauss (LG) beams serve as the exact solutions to the Helmholtz equation in free space. Consequently, when LG beams are employed for orbital angular momentum (OAM) multiplexing and atmospheric turbulence is neglected, the resulting field can be precisely represented by a superposition of LG modes, each carrying a distinct OAM value:

$$\varphi(x, y, z) = \sum_m a_m u_m(x, y, z) e^{jk_0 z} \quad (5.131)$$

where m denotes the m th order LG beam. The LG beams u_m fulfill the scalar Helmholtz equation in free space, i.e.

$$\nabla^2 (u_m(x, y, z) e^{jk_0 z}) + k_0^2 (u_m(x, y, z) e^{jk_0 z}) = 0 \quad (5.132)$$

and the beams are expressed by [88]

$$\begin{aligned} u_m(r, \phi, z) &= \sqrt{\frac{2p!}{\pi(p+|m|)!}} \frac{1}{w(z)} \left(\frac{r\sqrt{2}}{w(z)} \right)^{|m|} L_p^m \left(\frac{2r^2}{w^2(z)} \right) \exp \left(\frac{-r^2}{w^2(z)} \right) \\ &\exp \left(\frac{-jk_0 r^2 z}{2(z^2 + z_R^2)} \right) \exp \left(j(2p + |m| + 1) \tan^{-1} \left(\frac{z}{z_R} \right) \right) \exp(-jm\phi) \end{aligned} \quad (5.133)$$

where

$$\begin{aligned} w(z) &= w_0 \sqrt{1 + (z/z_R)^2} \\ z_R &= \frac{\pi w_0^2}{\lambda} \end{aligned} \quad (5.134)$$

where the initial beam waist is denoted by w_0 , the wavelength by λ , and the LG polynomial is specified with $p = 0$ and $L_{pm} = 1$ (implying the fundamental radial mode), the OAM modes are represented by ring-shaped beams, with the exception of

the $m = 0$ case which does not carry orbital angular momentum and may not exhibit a ring-like intensity profile [88].

It should be noted that these LG modes have been normalized and they are orthogonal to each other

$$\iint u_m^* u_n dS = \delta_{mn} \quad (5.135)$$

where $*$ stands for complex conjugate. It should be highlighted here that the normalization expressions employed in various works are slightly different. Specifically, the normalization used in the current work and in reference [88] differs from that in [5–7]. The normalization expression in [5–7] was tailored for the traditional coupled mode theory in optical waveguides and involves an additional factor $\frac{2\omega\mu}{\beta_m}$, where β_m is the propagation constant of the m th waveguide mode. However, when considering Laguerre–Gauss (LG) modes, the propagation constants of these modes are very close to each other. As a result, the additional factor in the normalization expression essentially becomes a constant for all LG modes. This means that the specific normalization used in Eq. (5.135) of the current work does not significantly alter the final results compared to the traditional coupled mode theory.

In scenarios involving turbulence, the coefficient a_m in Eq. (5.131) becomes dependent on the axial coordinate z . By assuming that the field are formed by the combination of the LG beams, one may combine Eqs. (5.130–5.132). Adopting the paraxial approximation for a_m and u_m , we have:

$$\sum 2 \frac{\partial a_m(z)}{\partial z} \left(\frac{\partial u_m}{\partial z} + jk_0 u_m \right) e^{jk_0 z} + 2\Delta n k_0^2 \sum a_m(z) u_m e^{jk_0 z} = 0 \quad (5.136)$$

The amplitude u_m is a slowly varying envelope comparing with the fast-oscillating phase term $e^{jk_0 z}$, and therefore, one has $\left| \frac{\partial u_m}{\partial z} e^{jk_0 z} \right| \ll \left| u_m \frac{\partial}{\partial z} (e^{jk_0 z}) \right|$. Hence, we can ignore $\frac{\partial u_m}{\partial z}$ since it is small compared to $jk_0 u_m$ and one has

$$\begin{aligned} \frac{da_m(z)}{dz} &= \sum_{m'} C_{mm'}(z) a_{m'}(z) \\ C_{mm'}(z) &= jk_0 \iint_{xy} u_m^* \Delta n u_{m'} dS \end{aligned} \quad (5.137)$$

It is crucial to acknowledge that while the zeroth radial order LG beams predominantly constitute the wave composition, they do not form a fully orthogonal basis, thereby implying the presence of a non-negligible fraction of other beam modes. These additional beam contributions can be interpreted as radiation losses inherent to the zeroth radial order LG beams. In our derivation, for the sake of simplicity, we have chosen to overlook these radiation losses while emphasizing that they can be subsequently incorporated into the final calculation of the coupled-power equations

(CPE), following the methodology employed by Marcuse in his sections, as a means of accounting for their impact [5–7].

The index fluctuation Δn causing the atmosphere turbulence has the following correlation function:

$$R(x - x', y - y', z - z') = \langle \Delta n(x, y, z) \Delta n(x', y', z') \rangle \quad (5.138)$$

The Fourier transform of above formula results in the Kolmogorov spectrum [88],

$$\begin{aligned} \mathbf{K}(k_x, k_y, k_z) &= \mathbf{FT}(R(x, y, z)) = (2\pi)^3 0.033 C_n^2 \left(k^2 + \frac{1}{L_0^2} \right)^{-11/6} f(k, k_l) \\ f(k, k_l) &= \exp\left(-\frac{k^2}{k_l^2}\right) \left(1 + 1.802 \left(\frac{k}{k_l} \right) - 0.254 \left(\frac{k}{k_l} \right)^{7/6} \right) \\ k^2 &= k_x^2 + k_y^2 + k_z^2 \\ k_l &= 3.3/l_0 \end{aligned} \quad (5.139)$$

where $\mathbf{FT}$ stands for the Fourier transform, with k_x , k_y and k_z representing the spatial frequencies in the x , y , and z directions respectively. The constants l_0 and L_0 are associated with the spectrum distribution, as specified in reference [88], and C_n^2 stands for the refractive-index structure parameter. It is worth noting that the expression presented in this work differs from that in [88] by a multiplicative factor of $(2\pi)^3$. This discrepancy arises due to the adoption of the general definition of the Fourier transform, as outlined in reference [93].

Given that the propagation distance z significantly exceeds the beam dimensions in the x and y directions, the underlying random process can be reasonably considered as two distinct and independent processes: one in the $x - y$ plane and the other along the z direction. Consequently, the correlation function can be approximately decomposed and expressed as a product of functions that individually characterize these two separate processes.

$$\begin{aligned} R(x, y, z) &\approx R_{xy}(x, y) f(z) \\ R_{xy}(x, y) &= \mathbf{FT}^{-1}(\mathbf{K}(k_x, k_y, 0)) \\ \langle f(z) f(z') \rangle &= \delta(z - z') \end{aligned} \quad (5.140)$$

when $\mathbf{FT}^{-1}$ stands for the inverse Fourier transform. We may express the power of each LG beam as:

$$P_m(z) = \langle a_m(z) a_m^*(z) \rangle \quad (5.141)$$

Using Eq. (5.137–5.141), we may convert Eq. (5.136) into the coupled power equation based on the procedures by Marcuse [5–7] or us [3]:

$$\frac{dP_m(z)}{dz} = \sum_{m'} \kappa_{mm'}(z)(P_{m'}(z) - P_m(z)) \quad (5.142)$$

The coupling coefficient is expressed as:

$$\begin{aligned} \kappa_{mm'}(z) &= k_0^2 \iint_{xy} \iint_{x'y'} u_m^*(x, y, z) u_m(x', y', z) u_{m'}(x, y, z) u_{m'}^*(x', y', z) R_{xy}(x - x', y - y') dS dS' \end{aligned} \quad (5.143)$$

Equation (5.142) has not considered any coupling between the LG beams and the radiation modes. By taking radiation modes into account, the exact coupled power equation should be

$$\frac{dP_m(z)}{dz} = \sum_{m'} \kappa_{mm'}(z)(P_{m'}(z) - P_m(z)) - \alpha_m(z)P_m(z) \quad (5.144)$$

where α_m stands for the radiation loss of each LG beam.

The coupling coefficient can be evaluated analytically as follows. Equation (5.143) can be rewritten as

$$\kappa_{mm'}(z) = k_0^2 \iint_{xy} (u_m^*(x, y, z) u_{m'}(x, y, z)) (R_{xy}(x, y) \otimes (u_m(x, y, z) u_{m'}^*(x, y, z))) dS \quad (5.145)$$

where $\otimes$ stands for two dimensional convolution. Equation (5.145) can be further simplified as:

$$\kappa_{mm'}(z) = \frac{1}{4\pi^2} k_0^2 \iint_{k_x k_y} |\mathbf{FT}(u_m^*(x, y, z) u_{m'}(x, y, z))|^2 \mathbf{K}(k_x, k_y, 0) dk_x dk_y \quad (5.146)$$

The radiation loss α_m can be estimated by accumulating (or integrating over) the coupling coefficients between the LG beams and the radiation modes. However, due to the intricacy in deriving the precise expressions for these radiation modes, we adopt an alternative approach by approximating the radiation modes with plane waves, a methodology commonly employed in the analysis of optical waveguides [10]. Hence, one may evaluate the loss as:

$$\begin{aligned} \alpha_m(z) &= \frac{1}{4\pi^2} k_0^2 \iint_{k_x k_y} \iint_{xy} \iint_{x'y'} u_m^*(x, y, z) u_m(x', y', z) e^{jk_x(x-x') + jk_y(y-y')} R_{xy}(x - x', y - y') dS dS' dk_x dk_y \\ &= k_0^2 \iint_{xy} \iint_{x'y'} u_m^*(x, y, z) u_m(x', y', z) \delta(x - x', y - y') R_{xy}(x - x', y - y') dS dS' \end{aligned}$$

$$\begin{aligned}
&= k_0^2 \iint_{xy} |u_m(x, y, z)|^2 R_{xy}(0, 0) dS \\
&= \frac{1}{4\pi^2} k_0^2 R_{xy}(0, 0) \\
&= \frac{1}{4\pi^2} k_0^2 \iint_{k_x k_y} \mathbf{K}(k_x, k_y, 0) dk_x dk_y
\end{aligned} \tag{5.147}$$

Strictly adhering to the principles, the plane wave approximation for radiation modes holds true primarily under the paraxial approximation during beam propagation, where k_x and k_y are considered small. Given that LG modes inherently possess dimensions significantly larger than the wavelength, the values of k_x and k_y are inherently small in comparison to the propagation constant k , thereby validating the paraxial approximation. Specifically, the upper bound for k_x and k_y is typically constrained by the grid resolution employed in numerical simulations, which is commonly on the order of 10^{-4} m. This resolution is sufficiently fine to accurately depict the variations within LG beams, with the largest k_x and k_y values being two orders of magnitude smaller than k . Consequently, Eq. (5.147) can be considered valid across the entire range of k_x and k_y values, allowing the integration intervals in Eq. (5.147) to be extended to $(-\infty, +\infty)$.

It is important to clarify that the radiation loss evaluated in Eq. (5.147) inherently includes the contributions stemming from the zeroth radial order LG modes themselves, as evidenced by their inclusion in Eq. (5.144). Therefore, to obtain the genuine radiation loss α_m that exclusively accounts for the radiation other than the zeroth radial order LG modes, it is necessary to subtract these related terms from the total loss calculated in Eq. (5.147), i.e. $\sum_{m'} \kappa_{mm'}(z)$, from Eq. (5.147).

Upon examining Eq. (5.147), a significant observation emerges: The loss experienced by each LG mode, inclusive of contributions from other LG modes, is solely influenced by atmospheric turbulence. This finding underscores that the beam parameters exclusively impact the coupling dynamics between LG modes, without directly altering their individual loss characteristics.

The CPE, represented by Eq. (5.144), can be efficiently solved using numerical integration techniques, such as the Runge–Kutta method. To facilitate the solution process and enhance clarity, it is advantageous to recast the CPE in a matrix form.

$$\frac{d\mathbf{P}(z)}{dz} = (\kappa(z) - \alpha(z))\mathbf{P}(z) \tag{5.148}$$

The solution can be written as:

$$\frac{d\mathbf{P}(z)}{dz} = \left(\lim_{\substack{N \rightarrow \infty \\ \Delta z \rightarrow 0}} \prod_{i=0}^N \exp((\kappa(z_i) - \alpha(z_i))\Delta z) \right) \mathbf{P}(0) \tag{5.149}$$

Given that the coupling coefficients exhibit minimal variation during propagation as long as the distance does not surpass the Rayleigh distance Z_R , Eq. (5.149) can

be further simplified as:

$$\begin{aligned}\kappa_{mm'}(z) &\approx \kappa_{mm'}(0) \\ \alpha_m(z) &\approx \alpha_m(0)\end{aligned}\tag{5.150}$$

Hence, we have the first order analytical approximate solution to Eq. (5.148):

$$\mathbf{P}(z) \approx \exp((\kappa(0) - \alpha(0))z)\mathbf{P}(0)\tag{5.151}$$

5.5.3 Results and Discussions

To validate the proposed coupled power model, a numerical example from [88] is adopted as a benchmark. This example utilizes turbulence parameters with $L_0 = 20\text{ m}$ for the outer scale and $l_0 = 5\text{ mm}$ for the inner scale. Additionally, the signal wavelength is set to be 850 nm, and the propagation distance is fixed at 1000 m. These parameters are identical to those specified in [88], ensuring a direct and fair comparison of results.

The numerical simulation parameters are outlined as follows: The simulation domain contains an area of $140\text{ cm} \times 140\text{ cm}$, discretized into a grid of 1024×1024 elements, marking an larger covering area compared to the configuration in [88], where a $70\text{ cm} \times 70\text{ cm}$ area was utilized with a 512×512 grid. Along the z-direction, random phase plates are inserted at every 50 m interval, each with uniquely generated phases by the approach adopted in [88]. The simulation employs the Split-Step Fourier Method (SSFM) to accurately model wave propagation, and for a comprehensive understanding of SSFM's workings, we recommend consulting [92].

To validate the SSFM Mont Carlo simulation outcomes, we initially verify them by computing the intensity and phase distributions of Orbital Angular Momentum (OAM) modes, specifically when the input is a pure OAM state with a topological charge of +5. The simulation parameters include a refractive index structure constant C_n^2 of $10^{-14}\text{ m}^{-2/3}$, an initial beam waist of 3 cm, and a propagation distance spanning 1000 m. One of the random realizations from the simulation is presented in Fig. 5.25, showing the intensity and phase distributions of the output field.

The figure demonstrates a close correspondence between the observed intensity/phase distribution and the results reported in [88], where the SSFM was likewise employed for the simulations. This validates the consistency and accuracy of our simulation approach.

Subsequently, we evaluate the scintillation index by adjusting the initial beam waist to 1.6 cm and maintaining a refractive-index structure parameter C_n^2 of $10^{-14}\text{ m}^{-2/3}$. Varying the propagation distance from 100 to 1000 m (corresponding

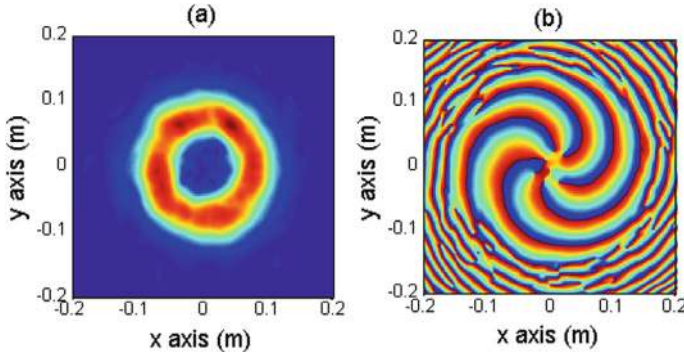


Fig. 5.25 Intensity (a) and phase (b) distribution of the output field when the input is OAM state +5. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

to a Rytov variance ranging from approximately 0.04 to 0.4), we assess the self-scintillation index of an input beam assumed to be in the OAM state 0. The results are plotted in Fig. 5.26, revealing a positive correlation between the scintillation index and the increasing Rytov variance. This trend aligns well with both theoretical predictions and experimental observations reported in [94]. Notably, at a propagation distance of 1000 m, the scintillation index reaches approximately 0.14, which is considered a reasonable value for a Gaussian beam as per the guidelines provided in [94]. Furthermore, [94] mentions that for a narrow collimated Gaussian beam with a Rytov variance of 0.4, the measured scintillation index is around 0.2, further substantiating our findings.

Our numerical model is able to explore the mode correlations, a crucial aspect for assessing MIMO channel capacity. For this analysis, we set the initial beam waist to be 1.6 cm, maintain a refractive-index structure parameter C_n^2 of $10^{-14} \text{ m}^{-2/3}$, and consider a propagation distance of 1000 m. We simulate the emission of OAM modes ranging from -9 to $+9$ at the input and their subsequent detection at the output. Assuming independent signals on each mode with equal power, which is a reasonable assumption during information transmission, we plot the correlations between OAM state 0 and other OAM states at the output in Fig. 5.27. The figure reveals that the strongest correlations exist between neighboring channels, with the correlation values ranging from 0.3 to 0.01. These values align with the expected range for channel correlations in the presence of atmospheric turbulence, as reported in [19].

The aforementioned calculations serve as a testament to the validity of our numerical model in simulating atmospheric turbulence effects. Utilizing the simulation parameters outlined, we proceed to calculate the final power transfer matrix, denoted as $\mathbf{M}$.

$$\mathbf{P}(z) = \mathbf{M}\mathbf{P}(0) \quad (5.152)$$

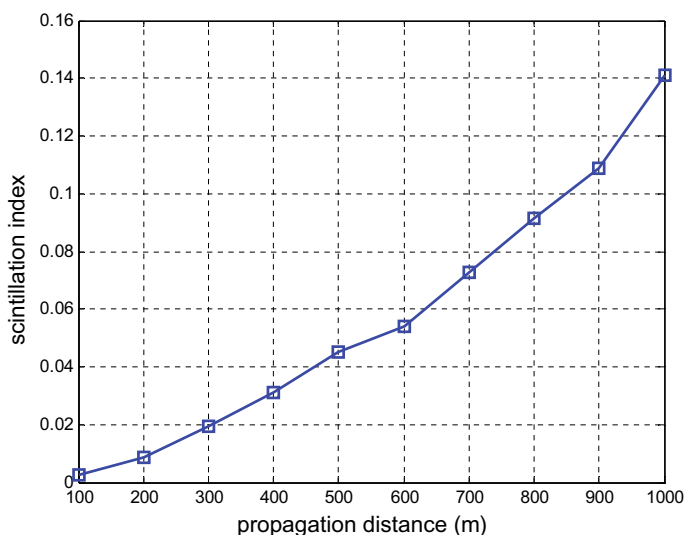


Fig. 5.26 Scintillation index versus propagation distance. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

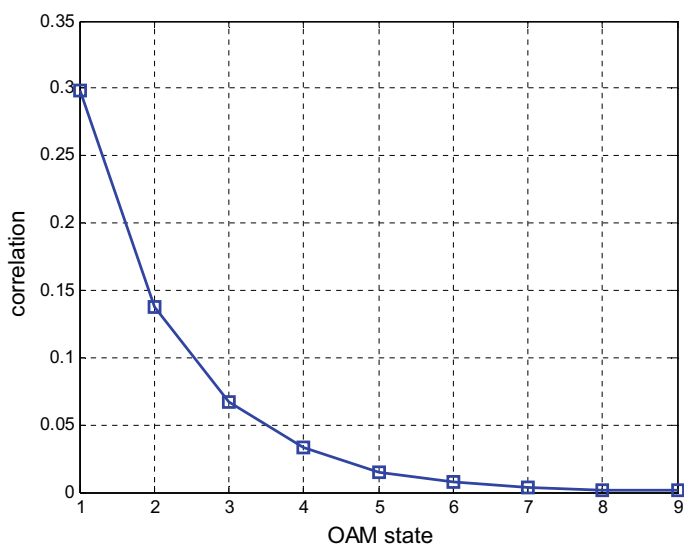


Fig. 5.27 Correlation between OAM state 0 and other OAM states when propagation distance $z = 1000$. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

It can either be calculated by the CPE model, i.e., Eqs. (5.149) and (5.151), or by the Mont-Carlo simulations.

The verification of our coupled mode power model commences with a comparative analysis between the data presented in [88] and the theoretical predictions generated by our models. Notably, the beams in [88] are LG beams with an initial beam waist of 1.6 cm, and the refractive-index structure parameter C_n^2 is set as $10^{-14} \text{ m}^{-2/3}$. It is noteworthy that when we adjust the simulation area and grid element count to match those specified in [88], our Mont-Carlo simulation results exhibit a high degree of similarity. The data in [88] serves as a benchmark for evaluating the accuracy of our proposed CPE, reinforcing its reliability.

The diagonal elements of the power transfer matrix represent the channel efficiency, signifying the proportion of power retained by each channel after propagation. Figure 5.28 illustrates the channel transmission efficiencies, where the results are derived from both Eqs. (5.149) and (5.151), as well as the data presented in [88]. As the mode order gets larger, the transmitted LG beams exhibit a diminishing capacity to retain power. The close alignment of the three curves underscores the precision of our formulas. Comparing the data from [88] with our solutions, the relative error remains below 8% for the “accurate” solution and within 13% for the approximate solution. When our Mont Carlo simulation results, utilizing an expanded simulation area, are factored in, the relative error is less than 4% for the “accurate” solution and less than 16% for the approximate solution, underscoring the robustness of our approach.

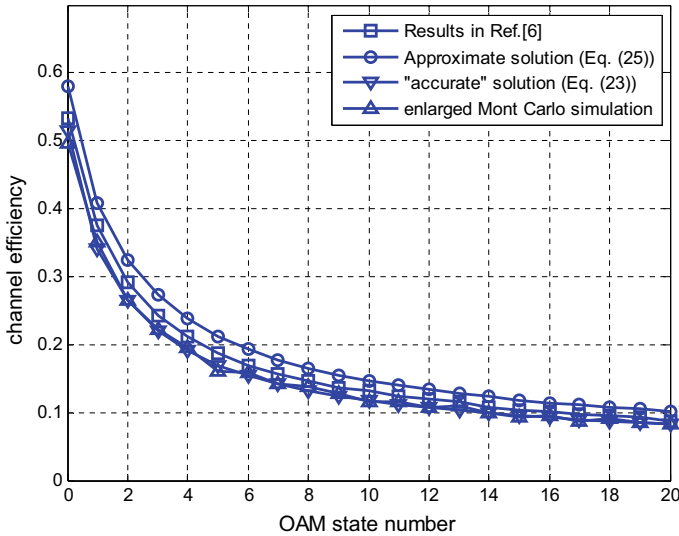


Fig. 5.28 OAM state conversion efficiency when the refractive index structure parameter C_n^2 is $10^{-14} \text{ m}^{-2/3}$ and the initial beam waist is 1.6 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

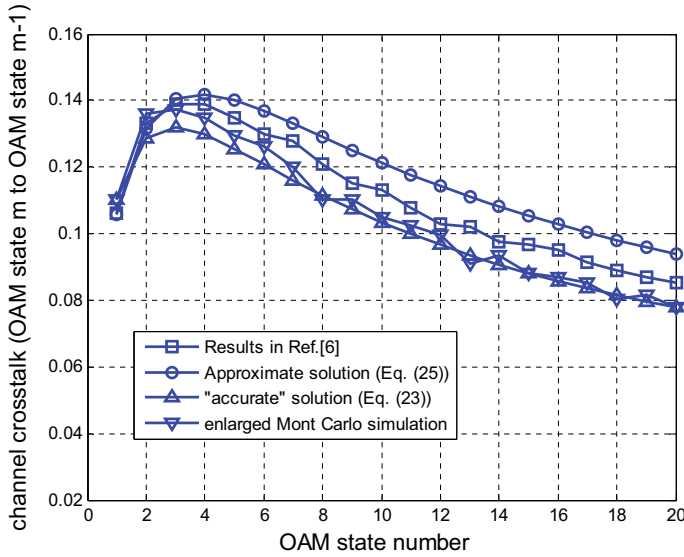


Fig. 5.29 OAM state crosstalk from OAM state m to OAM state $m-1$ when the refractive index structure parameter C_n^2 is $10^{-14} \text{ m}^{-2/3}$ and the initial beam waist is 1.6 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

Figures 5.29 and 5.30 depict the crosstalk coefficients between channels, represented by the off-diagonal elements of the transfer matrix M . Specifically, Fig. 5.29 shows the crosstalk from the OAM state m to the adjacent OAM state $m-1$, illustrating the interaction between neighboring channels. Conversely, Fig. 5.30 presents the crosstalk from the OAM state m to the OAM state $m-2$, highlighting a more distant crosstalk scenario. In both cases, the theoretical calculations, the outcomes of our Mont Carlo simulations, and the data from [88] exhibit a remarkable degree of consistency, validating the accuracy and reliability of our methods.

To verify the robustness of the proposed CPE, we employ a range of simulation parameters, including varied atmospheric turbulence intensities, diverse initial beam waist sizes, and different propagation distances. This comprehensive approach ensures that the CPE stands firm under diverse conditions, reinforcing its reliability and applicability.

Figures 5.31 and 5.32 illustrate the channel efficiency under varying atmospheric turbulence conditions. Specifically, Fig. 5.31 depicts the scenario with a refractive-index structure parameter C_n^2 of $10^{-15} \text{ m}^{-2/3}$, representing moderate turbulence, while Fig. 5.32 shows the case with $C_n^2 = 10^{-16} \text{ m}^{-2/3}$, indicating weak turbulence. In both cases, the initial beam waist is maintained at 1.6 cm. The results demonstrate remarkable accuracy across both moderate and weak turbulence environments. For moderate turbulence (Fig. 5.31), the relative error remains below 5% for the “accurate” solution and within 10% for the approximate solution. Notably, in

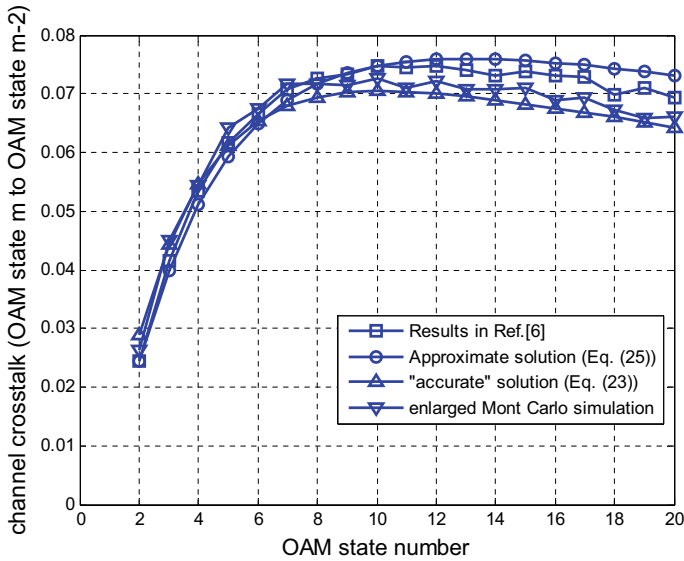


Fig. 5.30 OAM state crosstalk from OAM state m to OAM state $m-2$ when the refractive index structure parameter C_n^2 is $10^{-14} \text{ m}^{-2/3}$ and the initial beam waist is 1.6 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

the weaker turbulence scenario (Fig. 5.32), the relative error decreases to 0.06 and 0.6% respectively, further demonstrating the precision and reliability of our approach.

Figure 5.33 presents the channel efficiency analysis for a larger initial beam waist of 3 cm. To accommodate this larger waist, the numerical simulation area is expanded to $3 \text{ m} \times 3 \text{ m}$. This expansion is crucial because the Mont-Carlo simulation utilizes the Split-Step Fourier Method (SSFM) to simulate beam propagation, which assumes a periodic field distribution during the Fast Fourier Transform (FFT) process. Consequently, within the simulation region, plane waves continuously interact, potentially introducing a “reverse coupling” effect—a phenomenon where radiation modes could theoretically couple back into the LG beams, which does not occur in real LG beam propagation due to their rapid dissipation. This effect is evident in our previous figures, where we compared results from the enlarged simulation area with those from [88].

Figure 5.33 underscores that the results from the $3 \text{ m} \times 3 \text{ m}$ simulation area align more closely with our theoretical predictions. Additionally, it demonstrates that the CPE provides an excellent estimation of channel efficiency, with a relative error of less than 15% for both the “accurate” and approximate solutions. While further increasing the simulation area could potentially reduce this error, it would significantly escalate the computational costs.

Figure 5.34 shows the accuracy of channel efficiency as the propagation distance varies from 100 to 2000 m, focusing solely on channel 0 for simplicity. With an initial

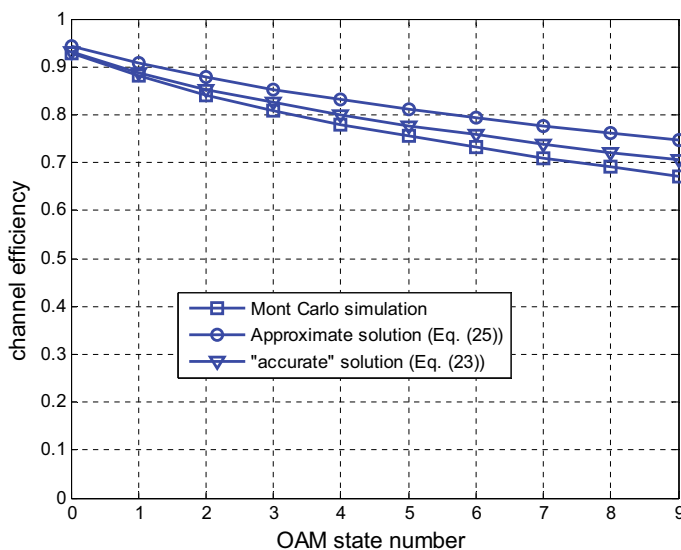


Fig. 5.31 OAM state conversion efficiency when the refractive index structure parameter C_n^2 is $10^{-15} \text{ m}^{-2/3}$, the initial beam waist is 1.6 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

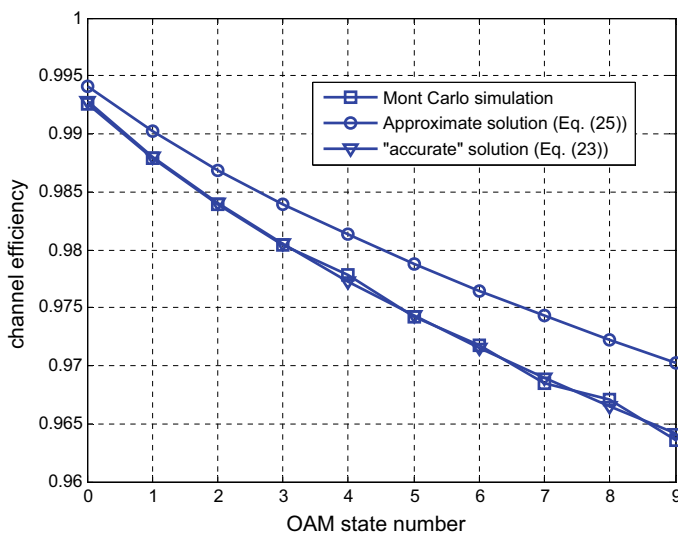


Fig. 5.32 OAM state conversion efficiency when the refractive index structure parameter C_n^2 is $10^{-16} \text{ m}^{-2/3}$, the initial beam waist is 1.6 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

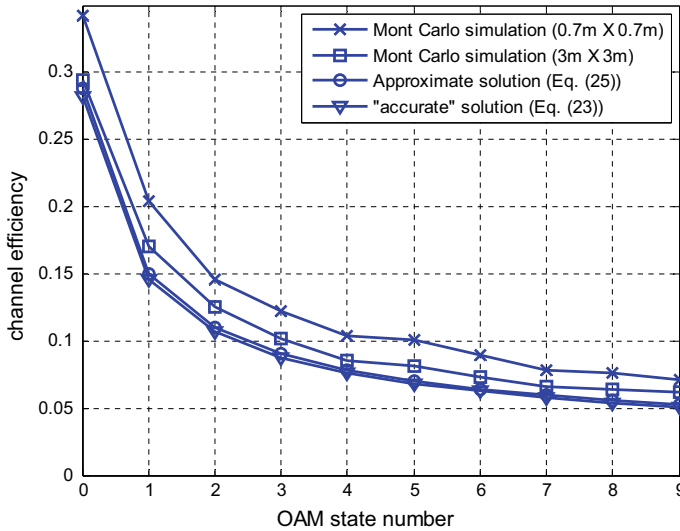


Fig. 5.33 OAM state conversion efficiency when the refractive index structure parameter C_n^2 is $10^{-14} \text{ m}^{-2/3}$ and the initial beam waist is 3 cm. The figure is reprinted/adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

beam waist of 1.6 cm and a turbulence strength of $C_n^2 = 10^{-14} \text{ m}^{-2/3}$, the Rayleigh range, as calculated from Eq. (5.134), is 946 m. To align with [88], the Mont Carlo simulation area remains 70 cm $\times$ 70 cm. The figure reveals that as the propagation distance increases, the relative error between the theoretical predictions and Mont Carlo simulations also rises. However, this error remains small until the propagation distance surpasses the Rayleigh range, both for the “accurate” solution and the approximate solution. Notably, even when the propagation distance significantly exceeds the Rayleigh range, the “accurate” solution maintains a relative error of approximately 10% compared to the Mont Carlo simulations. The primary source of this error escalation is the computational error arising from beam broadening, which can be mitigated by enlarging the simulation area in the Mont Carlo simulations.

As is evident from Figs. 5.28, 5.29, 5.30, 5.31, 5.32, 5.33, and 5.34, both the approximate and “accurate” solutions of the CPE exhibit high accuracy when compared to Mont Carlo simulation results within the Rayleigh range. This concordance arises from the fact that the coupling coefficients vary little during propagation distances shorter than the Rayleigh distance Z_R . Consequently, from a practical perspective, the first-order approximate solution is precise to predict the propagation characteristics of OAM states in a wide range of applications. The minor deviations between the theoretical “accurate” solution and Mont Carlo simulations primarily arise due to the “reverse coupling” effect, which is an artifact of the limited simulation area in Mont Carlo simulations. Despite optimizing the numerical parameters from [88] in some of our simulations, the “reverse coupling” effect persists, particularly under strong turbulence and long propagation distances. This effect intensifies

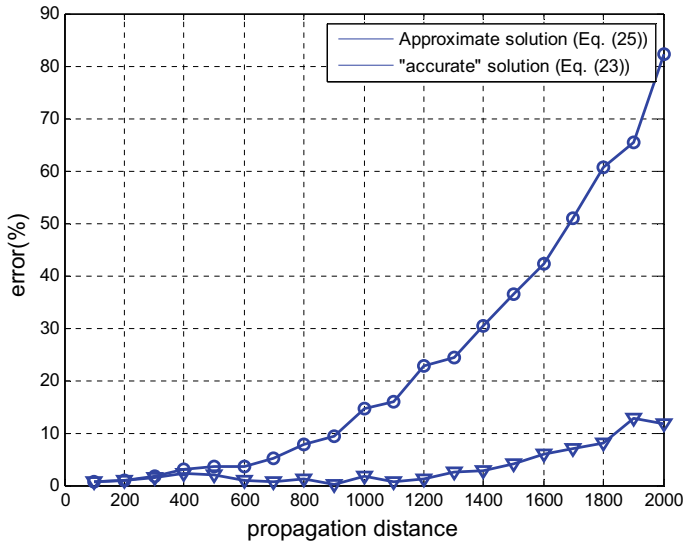


Fig. 5.34 OAM state 0 conversion efficiency deviation between the Mont Carlo simulation and the theory versus propagation distance, the refractive index structure parameter C_n^2 is $10^{-14} \text{ m}^{-2/3}$ and the initial beam waist is 1.6 cm. The simulation area is $70 \text{ cm} \times 70 \text{ cm}$. The figure is reprinted/ adapted with permission from [Opt. Express 25, 31964–31976 (2015)] © The Optical Society of America [82]

beyond the Rayleigh range as the beam spreads significantly, enhancing the likelihood of “reverse coupling.” Generally, the extent of “reverse coupling” is influenced by beam size, turbulence strength, and propagation distance. While increasing the simulation area can mitigate this effect, it comes at the cost of significantly higher computational demands.

5.5.4 Summary

In summary, we have formulated an efficient theoretical framework to analyze the influence of atmospheric turbulence on the propagation of OAM modes in free space. This theory is based on the coupled mode theory, from which we derived a coupled average power equation, incorporating analytical expressions for coupling coefficients and modal losses. The proposed CPE can be solved numerically, offering a cost-effective alternative to Monte Carlo simulations, or approximated using its first-order solution for further simplification. The validity of our CPE results has been confirmed by comparison with published data and Monte Carlo simulations. By combining the approximate analytical solution with the coupled average power equation, we enable highly efficient predictions of the dynamic behavior of OAM modes in atmospheric turbulence.

5.6 Capacity Analysis for Free Space Coherent Optical MIMO Transmission Systems: With and Without Adaptive Optics

Part of the section reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

The random propagation in the free space is analyzed continually in this section. Based on the CMT, it is found that the channel matrix of free space modes can be obtained analytically. Therefore, it paves the way to analyze the capacity of the free space coherent optical MIMO transmission system. In this section, two scenarios are examined: one with adaptive optics compensation and the other without. It is widely acknowledged that adaptive optics compensation remarkably enhances system performance, particularly when MIMO algorithms are absent. However, this generalization may not hold true for coherent MIMO systems. Our findings indicate that, under weak to moderate turbulence conditions, utilizing a phase-only wavefront corrector in such systems increase the mean eigenvalue of the coherent system capacity matrix (denoted as $\mathbf{H}\mathbf{H}^H$), resulting in a more uniform eigenvalue distribution. Specifically, it diminishes the maximum eigenvalue while elevating the average, thus enhancing system capacity in the absence of channel information, as sub-channels are equally powered. On the other hand, when channel information is fully accessible and the water-filling algorithm is implemented to optimize the power allocation, the system with adaptive optics may exhibit reduced performance, as capacity becomes more reliant on large eigenvalues, especially at low signal-to-noise ratios (SNRs). In the scenarios of intense turbulence, adaptive optics reduces both the mean and maximum eigenvalues of the capacity matrix $\mathbf{H}\mathbf{H}^H$, leading to a decline in system capacity, regardless of the availability of channel information.

5.6.1 Introduction

Space division multiplexing (SDM) has been proposed to further boost the transmission capacity of the optical communication systems [83]. In Free Space Optical (FSO) transmission, SDM is achieved by encoding signals onto a collection of spatial modes. These spatial modes, such as the Laguerre–Gauss (LG) modes [96], the Bessel-Gauss modes [97] or any other spatial modes under the condition that they maintain orthogonal to each other during the free space propagation.

The orthogonality of modes, notably the Laguerre-Gaussian (LG) modes, is impeccably preserved in isotropic and homogeneous media. However, the presence of atmospheric turbulence [19, 82, 88, 91, 98–101], arising from factors such as rain, water vapor, and smog, introduces distortions. This turbulence results in random scattering, which distorts the transmission beams/channels and can lead to crosstalk among modes and consequently deteriorate system capacity [99–101]. In response, numerous researchers have explored techniques to mitigate the effects of atmospheric

turbulence. One effective strategy involves utilizing phase-only adaptive optics [102–104]. This approach involves transmitting a beacon wave, which serves as a monitor for phase errors in the transmission channel. The detected phase errors are then compensated for by applying a phase-only spatial mask. The channel monitoring can be implemented by either sending the beacon wave in the opposite direction to the signal [102] or by utilizing a beam traveling in the same direction but at a distinct wavelength [103, 104].

For the non-coherent FSO system employing intensity channels, the capacity has been extensively researched [105–108]. For the coherent MIMO system where the input signals are not constrained to be real and non-negative, the system capacity is dictated by the channel capacity matrix $\mathbf{H}\mathbf{H}^H$ [96]. Here, not only the diagonal elements of $\mathbf{H}$, which can be enhanced through adaptive optics, are crucial, but also the off-diagonal elements, representing crosstalk between modes, play a pivotal role. Recent studies have focused on the system capacity of FSO communication systems in the presence of crosstalk. In the free space non-MIMO SDM systems [88, 102–104], the capacity/performance has been discussed with [102–104] and without [88] the implementation of phase-only adaptive optics. N. Zhao et al. conducted a comprehensive analysis on the capacity of free-space coherent MIMO systems in [96], where crosstalk primarily arisen from the finite pupil size. Additionally, further investigations into free-space MIMO systems, both with and without the use of non-phase-only adaptive optics (also known as “full wave adaptive optics”), have been presented in [99, 100]. However, the conclusion can not be directly applied to the phase-only adaptive optics based FSO systems. Furthermore, [99, 100] dealt with the quantum–mechanical (Holevo-limit) capacity, which differed from the Shannon-limit capacity commonly considered in coherent optical communication systems. Our recent work [101] analyzed the Shannon-limit capacity upper bound for coherent MIMO systems without adaptive optics under atmospheric turbulence and identified optimal mode sets. However, the study in [101] did not account for scenarios where full channel information was available or adaptive optics were employed.

Despite the pioneering studies [88, 96, 99, 100, 102–108] examining system capacity in free-space SDM systems, there remains a gap in investigating the capacity of coherent optical MIMO systems under atmospheric turbulence when leveraging currently available phase-only adaptive optics. Additionally, a comparative analysis of the turbulence-affected coherent MIMO system capacity, both with and without channel information, has yet to be conducted.

In this section, we conduct a comprehensive examination of the system capacity for free-space coherent MIMO systems, considering both scenarios with and without the incorporation of phase-only adaptive optics. Our objectives are to address the following questions: How does the performance of coherent MIMO systems differ when phase-only adaptive optics are employed versus when they are absent? What impact does the availability of channel information have on this performance? Furthermore, how does the system capacity respond as turbulence strength varies from weak/moderate to strong conditions? Lastly, we aim to explore the possibility that, under certain scenarios, the use of phase-only adaptive optics might inadvertently lead to a degradation in system capacity.

5.6.2 Theory

5.6.2.1 The LG Modes Propagation in Free Space

Without loss of generality, this work assumes the use of Laguerre-Gaussian (LG) modes, as they constitute a complete orthonormal basis in free space and are frequently employed in free-space spatial division multiplexing (SDM) systems. These modes are characterized by the initial beam waist w_0 , the free-space wavelength λ_0 , the azimuth order m , and the radial order p . They can be mathematically represented as follows [88, 101]:

$$u_m^p(r, \phi, z) = \sqrt{\frac{2p!}{\pi(p+|m|)!}} \frac{1}{w(z)} \left(\frac{r\sqrt{2}}{w(z)} \right)^{|m|} \exp\left(\frac{-r^2}{w^2(z)}\right) L_p^{|m|}\left(\frac{2r^2}{w^2(z)}\right) \exp\left(\frac{-jk_0 r^2 z}{2(z^2 + z_R^2)}\right) \exp\left(j(2p + |m| + 1) \tan^{-1}\left(\frac{z}{z_R}\right)\right) \exp(-jm\phi) \quad (5.153)$$

with

$$w(z) = w_0 \sqrt{1 + (z/z_R)^2} \quad (5.154)$$

$$z_R = \frac{\pi w_0^2}{\lambda_0}$$

where r , ϕ , and z represent the radius, angle, and propagation distance respectively in the cylindrical coordinate system. The initial beam waist is denoted by w_0 , and λ_0 is the free-space wavelength. The indices m and p correspond to the azimuth order and radial order of the associated Laguerre polynomial L_{pm} . In the absence of atmospheric turbulence, the propagation of LG modes can be formulated by Eqs. (5.153–5.154) while they maintain orthogonal to each other.

Following the formulation in the previous section, the wave under turbulence should fulfill [92, 101]:

$$\nabla^2 \varphi + k_0^2 (1 + 2\Delta n) \varphi = 0 \quad (5.155)$$

where k_0 is $2\pi/\lambda_0$, Δn the zero-mean refractive index variation. The correlation function of the index variation should be:

$$R(x - x', y - y', z - z') = \langle \Delta n(x, y, z) \Delta n(x', y', z') \rangle \quad (5.156)$$

As is stated in the previous section, the Fourier transform of the above Eq. (5.160) is the Kolmogorov spectrum or other forms of spectrum functions. Without loss of generality, we adopt the Kolmogorov spectrum [88, 101]:

$$\begin{aligned}
\mathbf{K}(k_x, k_y, k_z) &= \mathbf{FT}(R(x, y, z)) = (2\pi)^3 0.033 C_n^2 \left(k^2 + \frac{1}{L_0^2} \right)^{-11/6} f(k, k_l) \\
f(k, k_l) &= \exp\left(-\frac{k^2}{k_l^2}\right) \left(1 + 1.802 \left(\frac{k}{k_l} \right) - 0.254 \left(\frac{k}{k_l} \right)^{\frac{7}{6}} \right) \\
k^2 &= k_x^2 + k_y^2 + k_z^2 \\
k_l &= 3.3/l_0
\end{aligned} \tag{5.157}$$

where $\mathbf{FT}$ denotes the Fourier transform, while k_x , k_y , and k_z represent the corresponding spatial frequencies. Furthermore, l_0 , L_0 , and C_n^2 are parameters that are associated with the shape of the spectrum and the strength of the turbulence, respectively [88, 101].

In our simulation, the propagation distance z is divided into several sections. For each of these sections, a random phase screen is introduced. These phase screens are generated by passing a two-dimensional Gaussian white noise through a filter, where the frequency domain response of this filter is the square root of $\mathbf{K}(k_x, k_y, 0)$. This multi-phase screen model is extensively employed in simulating light propagation through atmospheric turbulence [19, 82, 88, 98, 104].

5.6.2.2 Channel Model and Capacity Calculation

In the system under consideration, a limited number of LG modes, specifically Q LG modes, are employed for transmission. Without loss of generality, these Q LG modes are assumed to differ either in their radial order p or their azimuth order m within the coherent MIMO (Multiple-Input Multiple-Output) system. This implies that both the number of transmitters and receivers are equal to Q . The received signal vector $\mathbf{r}$ can be calculated from the transmitted signal vector $\mathbf{x}$ through [96, 101]

$$\mathbf{r} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{5.158}$$

Here, $\mathbf{H}$ represents the channel matrix, where its diagonal elements signify the residual amplitudes of the transmitted modes, while the off-diagonal elements indicate the crosstalk among these modes. Additionally, $\mathbf{n}$ denotes the Gaussian independent and identically distributed (i.i.d.) noise vector introduced at both the transmitter and the receiver. In contrast to free-space intensity channels [105–108], the aforementioned channel model imposes no real and non-negative constraints on the input vector $\mathbf{x}$, aligning it with the coherent MIMO channel model presented in [96]. The capacity of such a system can be formulated by [96]

$$C = \sum_{q=1}^Q \log_2 \left(1 + \lambda_q \frac{P_q}{\sigma^2} \right) \tag{5.159}$$

where λ_q denotes the eigen value of $\mathbf{H}^H\mathbf{H}$, P_q represents the power on the q th sub-channel which is related to the q th eigen vector of $\mathbf{H}^H\mathbf{H}$, and σ^2 is computed by $\langle n_q(t)n_q^*(t) \rangle$ which denotes the noise power on each sub-channel.

Two scenarios are usually taken into account while computing Eq. (5.159). When the channel matrix $\mathbf{H}$ is unknown at the transmitter side, an equal power allocation strategy for each channel is typically assumed to maximize the system capacity.

$$P_q = \frac{P_{tot}}{Q} \quad (5.160)$$

where P_{tot} is for the total signal power. In scenarios where pilot signals can be transmitted to estimate the channel matrix $\mathbf{H}$, the water-filling algorithm can be employed to optimize the power allocation on each sub-channel [109].

$$\begin{aligned} P_q &= \left(\mu - \frac{\sigma^2}{\lambda_q} \right)^+ \\ P_{tot} &= \sum_q P_q \\ (x)^+ &= \begin{cases} 0 & (x < 0) \\ x & (x > 0) \end{cases} \end{aligned} \quad (5.161)$$

where μ is a constant for all of the sub-channels. Since the free-space channel experiences variations on the order of milliseconds, channel measurement becomes feasible for free-space optical transmission systems operating at data rates exceeding Gbps [102].

5.6.2.3 Analysis of the Mean Eigen Value of the Capacity Matrix $\mathbf{H}^H\mathbf{H}$ with and Without Adaptive Optics

We define the transfer coefficient from the i -th LG mode to the j -th LG mode as H_{ij} , where i and j correspond to the indices of the respective LG modes. It is important to clarify that H_{ij} may not necessarily be an element of the channel matrix $\mathbf{H}$, as $\mathbf{H}$ only includes H_{ij} as its element if both mode i and mode j belong to the set of Q modes (i.e., $i \leq Q$ and $j \leq Q$) that have been selected for transmission. To distinguish between scenarios with and without adaptive optics, we will use the superscripts ‘ad’ or ‘non’ respectively.

First of all, it is easy to derive that the following inequality stands under the weak/moderate turbulence:

$$\langle |H_{ii}^{ad}|^2 \rangle > \langle |H_{ii}^{non}|^2 \rangle \quad (5.162)$$

where $\langle \rangle$ stands for expectation. It is valid since adaptive optics increases the receiving power by compensating the wave-front distortion, which has been widely proven by the experimental and numerical studies in [102–105].

It is straight forward that

$$\sum_{i=1}^Q \lambda_i = \text{Tr}(\mathbf{H}^H \mathbf{H}) = \sum_{i=1}^Q \sum_{j=1}^Q |H_{ij}|^2 \quad (5.163)$$

where tr is the symbol for the trace of the matrix. Therefore, the average of the eigen values of $\mathbf{H}^H \mathbf{H}$ is decided by the diagonal element summation.

Due to the energy conservation law and the completeness and orthonormality of the LG mode basis, one has

$$\sum_{j=1}^{\infty} |H_{ij}^{ad}|^2 = \sum_{j=1}^{\infty} |H_{ij}^{non}|^2 = 1 \quad (5.164)$$

It is widely acknowledged and experimentally/numerically verified that power transfer predominantly occurs between neighboring LG modes [19, 82, 88, 98], specifically those with close values of the azimuth order m and radial order p . Consequently, for the “middle” modes with a considerable number of neighboring modes included in the selected Q -mode set, one has the approximate equality under the weak/moderate turbulence condition:

$$\left\langle \sum_{j=1}^Q |H_{ij}^{ad}|^2 \right\rangle \approx \left\langle \sum_{j=1}^Q |H_{ij}^{non}|^2 \right\rangle \approx 1 \quad (5.165)$$

where i stands for a specific mode number ‘in the middle’.

Equation (5.162) indicates that the implementation of adaptive optics results in a greater concentration of power on the central mode i . For modes with a limited number of neighboring modes within the selected Q -mode set, the summation of the squares of the row elements of $\mathbf{H}$ primarily depends on the value of $\langle |H_{ii}|^2 \rangle$, and one has

$$\left\langle \sum_{j=1}^Q |H_{ij}^{ad}|^2 \right\rangle > \left\langle \sum_{j=1}^Q |H_{ij}^{non}|^2 \right\rangle \quad (5.166)$$

A mode varies between “the mode with quite a few neighboring modes” and “the mode with less neighboring modes” should have $\left\langle \sum_{j=1}^Q |H_{ij}^{ad}|^2 \right\rangle$ to be either larger or close to $\left\langle \sum_{j=1}^Q |H_{ij}^{non}|^2 \right\rangle$. Regarding this fact, one has the inequality under the weak/

moderate turbulence assumption

$$\begin{aligned} \left\langle \sum_{i=1}^Q \sum_{j=1}^Q |H_{ij}^{ad}|^2 \right\rangle &> \left\langle \sum_{i=1}^Q \sum_{j=1}^Q |H_{ij}^{non}|^2 \right\rangle \\ \frac{\left\langle \sum_{i=1}^Q \lambda_i^{ad} \right\rangle}{Q} &> \frac{\left\langle \sum_{i=1}^Q \lambda_i^{non} \right\rangle}{Q} \end{aligned} \quad (5.167)$$

Equation (5.167) suggests that under weak to moderate turbulence conditions, the summation of the diagonal elements or the mean eigenvalue of the capacity matrix is enhanced by the application of adaptive optics. Consequently, a superior system capacity performance is anticipated when channel information is unavailable and equal power allocation is assumed over all sub-channels.

5.6.3 Results and Discussions

Numerical simulations have been conducted to evaluate the system's performance with the same parameters as specified in [88, 101]. The signal operates at a wavelength of 850 nm, and the turbulence spectral parameters are set to $L_0 = 20$ m and $l_0 = 5$ mm. The transmission distance of the system is 1000 m, with random phase screens inserted every 50 m along the propagation path [88, 101]. Monte Carlo simulations with 1000 realizations were executed to ensure a balance between the accuracy of the averaged results and computational efficiency, as a smaller number of realizations can lead to larger variances in the outcomes.

The simulation area for characterizing beam propagation spans a $280 \text{ cm} \times 280 \text{ cm}$ square, discretized into a grid of 2048×2048 elements. The initial beam waist w_0 of the LG modes is set to 1.6 cm. A total of 76 LG modes, with radial order p ranging from 0 to 3 and azimuth order m from -9 to $+9$, are considered for transmission, and they propagate collinearly. It is noteworthy that while selecting a different mode set may alter the system's capacity value, the behaviors with and without adaptive optics will remain qualitatively similar. According to [96], the root means that square (RMS) waist size of the LG modes can be computed by $w_z \sqrt{2p + |l| + 1}$, revealing that the largest RMS waist sizes at $z = 0$ and $z = L$ are $4w_0$ and $4w_L$, respectively. Given that the impact of finite pupil size has been thoroughly analyzed in [96], we assume that the lenses at the transmitter and receiver are sufficiently large to capture most of the energy from the 76 LG modes. This is achieved by ensuring their pupil radii significantly exceed the largest beam waist, $4w_0$ and $4w_L$, in our scenario. This configuration allows us to disregard the pupil size effects and concentrate solely on the influence of atmospheric turbulence.

As is previously discussed, the simulation includes two distinct scenarios: one with adaptive optics and the other without. In the absence of adaptive optics, the

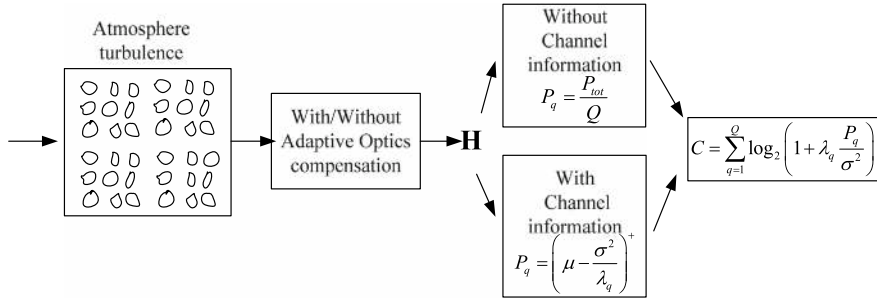


Fig. 5.35 The schematic to evaluate the capacity during the simulation. This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

beam transmits through the atmosphere turbulence and correlates with the ideal LG modes at the receiver to derive the corresponding transmission matrix $\mathbf{H}$. When adaptive optics is employed, a beacon beam with radial order 0 and azimuth order 0 is transmitted to monitor the wavefront. This beacon beam can operate at the same or a different wavelength from the signal [102–104]. Here, a beacon wavelength of 852 nm is chosen, which is close to the 850 nm signal wavelength to avoid the impact of the wavelength difference on the capacity. It has been verified that the wavelength separation between the signal and the beacon wave can be as large as 10 nm without significantly affecting the capacity. The beacon beam is designed to have a beam waist equivalent to that of the largest LG mode and is detected at the receiver. A phase-only spatial mask array is utilized to compensate for the phase distortions incurred by the 76 transmitted LG modes. It is assumed that sufficient phase-adjustment mirrors are available to fully rectify the phase front distortions detected by the beacon beam. This assumption facilitates an investigation into the theoretical capacity limit of the free-space transmission system when compensated by the adaptive optics. Figure 5.35 illustrates the schematic for determining the system capacity under these conditions.

5.6.3.1 Verification of the Equations in Sect. 5.2.3

As a demonstration for Sect. 5.2.3, an illustrative scenario involving weak to moderate turbulence is presented, with C_n^2 set to $10^{-15} \text{ m}^{-2/3}$. To ensure statistical significance, a total of 1000 realizations are simulated to calculate the average diagonal elements of the capacity matrix $\mathbf{H}^H \mathbf{H}$. Given that the simulation implements LG modes with azimuth order m ranging from -9 to $+9$ and radial order p from 0 to 3, the lowest-order LG mode (i.e., $m = 0, p = 0$) is expected to have the highest number of neighboring modes. The resulting average diagonal elements for the adaptive optics corrected case ($\mathbf{H}^{\text{ad}H} \mathbf{H}^{\text{ad}}$) and the non-AO case ($\mathbf{H}^{\text{non}H} \mathbf{H}^{\text{non}}$) are 0.9967 and 0.9969, respectively. These values are quite close to 1, validating Eq. (5.165). Similar trends are observed for diagonal elements corresponding to the modes with a large number

of neighboring modes. The average differences between the corresponding diagonal elements are less than 0.004.

On the other hand, the LG mode with $m = 9$ and $p = 3$ has the least neighboring modes, which results in the corresponding diagonal elements average values as 0.8427 and 0.6968 for the cases with and without adaptive optics, thus verifying Eq. (5.166).

Combining Eqs. (5.165 and 5.166), one may verify Eq. (5.167). The summation of the diagonal elements for $\mathbf{H}^{adH} \mathbf{H}^{ad}$ and $\mathbf{H}^{nonH} \mathbf{H}^{non}$ are 72.8994 and 71.3498 in average respectively.

5.6.3.2 System Capacity Analysis

In the simulation for the MIMO system, the turbulence strength C_n^2 changes and it is initially set to be $10^{-14} \text{ m}^{-2/3}$, which represents the moderate turbulence strength. The capacity calculations, averaged over 1000 realizations, are presented in Fig. 5.36. Since the two curves for the results with and without adaptive optics are quite close to each other, we have the subplots inside the figures to magnify the deviations.

These subplots reveal distinct behaviors in system capacity based on the availability of channel information. In the absence of channel information, a uniform power allocation across sub-channels is assumed, and the system equipped with

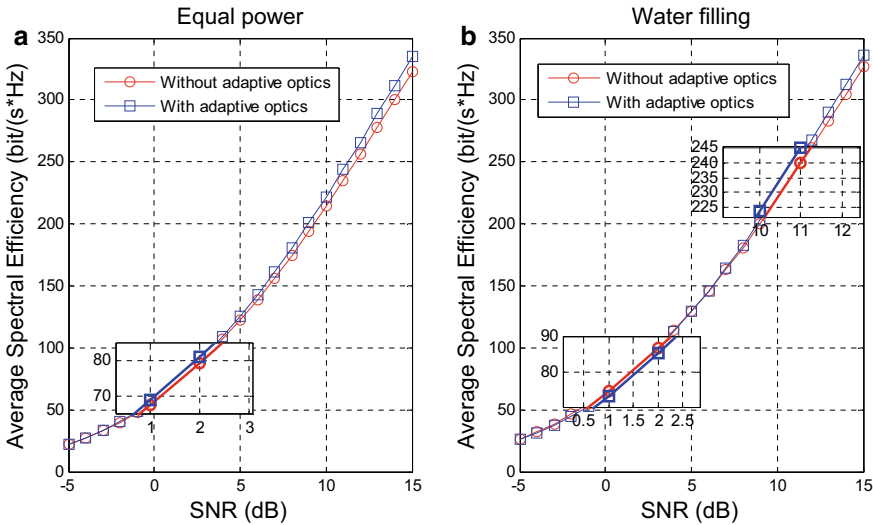


Fig. 5.36 The capacity of the 76 modes transmission system ($C_n^2 = 10^{-14} \text{ m}^{-2/3}$). **a** Without the channel information and equal power is assumed for each sub-channel. **b** With the channel information and the watering filling algorithm is implemented to optimize the power allocation. This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

adaptive optics exhibits enhanced performance. When channel information is accessible and the water-filling algorithm is employed to optimize the power allocation, an interesting trend emerges: at low signal-to-noise ratios (SNRs), the system without adaptive optics achieves higher capacity, whereas the system with adaptive optics surpasses it in the high SNR regime. Notably, the threshold SNR for this performance transition is identified as 5 dB.

To elucidate the behaviors observed in Fig. 5.36, Fig. 5.37 depicts the eigenvalue distribution of the capacity matrix $\mathbf{H}\mathbf{H}^H$, sorted in descending order and averaged across 1000 realizations. Consistent with the summaries in Sects. 5.2.3 and 5.3.1, it is evident that the mean eigenvalue is higher for the system utilizing adaptive optics. This explains the improved system capacity when equal power is allocated to each sub-channel, as the system with adaptive optics possesses higher eigenvalues.

Furthermore, the application of adaptive optics tends to amplify smaller eigenvalues while attenuating larger ones. This characteristic explains why, in the low signal-to-noise ratio (SNR) regime where the water-filling algorithm is employed, the adaptive optics system might experience a capacity degradation. Specifically, the water-filling algorithm preferentially allocates power to sub-channels with larger eigenvalues, particularly in the low SNR domain. Consequently, the redistribution of eigenvalues by adaptive optics can lead to a degraded overall system capacity.

When C_n^2 is $5 \times 10^{-15} \text{ m}^{-2/3}$, indicating a relatively weaker turbulence environment, the corresponding capacity and eigenvalue analyses are presented in Figs. 5.38 and 5.39, respectively. These results mirror the previous findings, leading to similar conclusions. Specifically, the incorporation of adaptive optics enhances performance

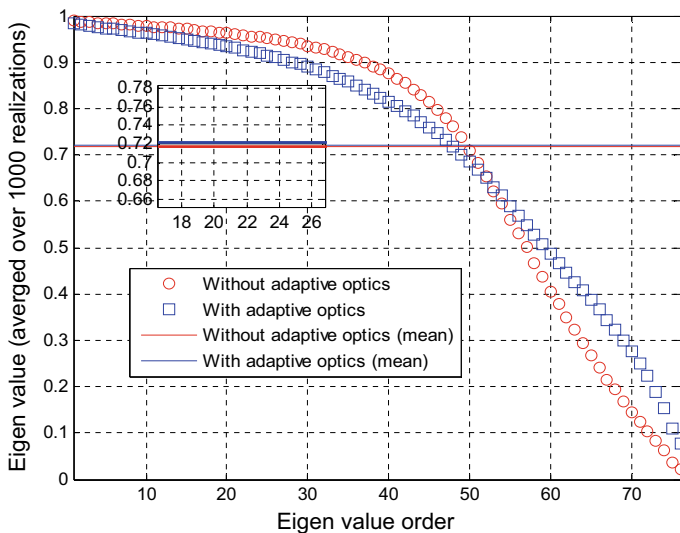


Fig. 5.37 The eigen value distribution of the capacity matrix $\mathbf{H}\mathbf{H}^H$ ($C_n^2 = 10^{-14} \text{ m}^{-2/3}$). This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

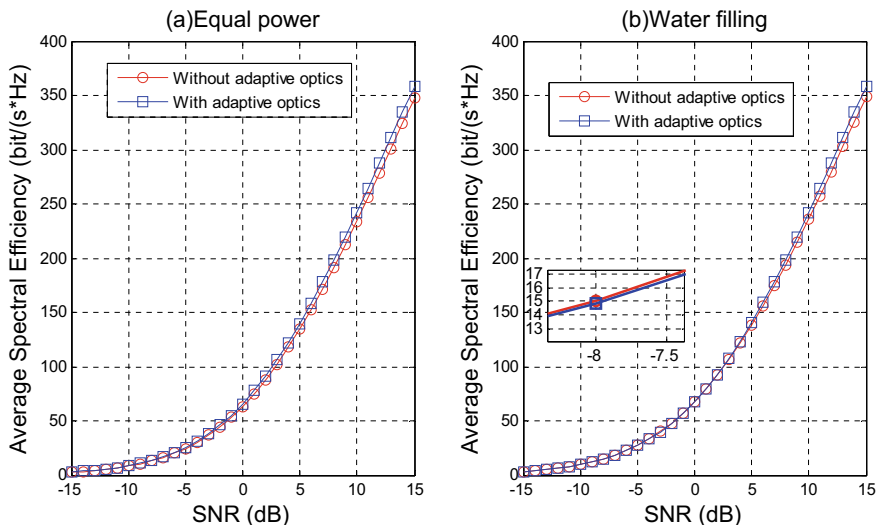


Fig. 5.38 The capacity of the 76 modes transmission system ($C_n^2 = 5 \times 10^{-15} \text{ m}^{-2/3}$). **a** Without the channel information and equal power is assumed for each sub-channel. **b** With the channel information and the watering filling algorithm is implemented to optimize the power allocation. This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

in two scenarios: firstly, when channel information is unavailable and power is evenly distributed among sub-channels; secondly, in the high signal-to-noise ratio (SNR) regime when channel information is available and power allocation is optimized using the water-filling algorithm.

However, a notable difference emerges in the transition threshold SNR. Compared to the previous case, this threshold shifts from 5 to 0 dB. Below this new threshold, the system without adaptive optics outperforms the adaptive optics equipped system when the water-filling algorithm is applied with known channel information. This phenomenon arises from the close values of the largest eigenvalues in both systems under weaker turbulence conditions. Consequently, the adaptive optics induced capacity degradation occurs at a lower SNR threshold.

As C_n^2 escalates to $10^{-13} \text{ m}^{-2/3}$, signifying a regime of strong turbulence, the summary for the mean eigenvalue of the capacity matrix in Sect. 5.2.3 becomes inapplicable. The results presented in Fig. 5.40 reveal a detrimental effect of adaptive optics compensation on system performance, irrespective of whether channel information is available or not.

Further analysis of the eigenvalue distribution in Fig. 5.41 sheds light on this phenomenon. Despite the enhancement of the diagonal elements of the transfer matrix $\mathbf{H}$ through adaptive optics compensation, the eigenvalues are observed to decrease. This counterintuitive outcome suggests that under conditions of strong turbulence, the corrective measures introduced by adaptive optics may inadvertently

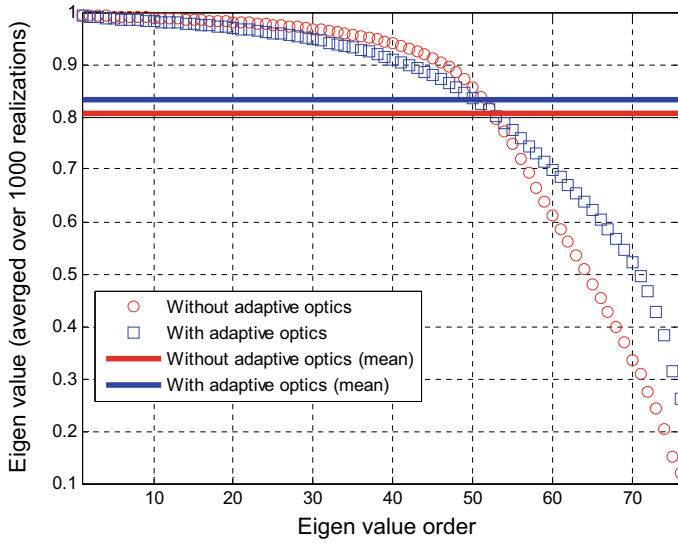


Fig. 5.39 The eigen value distribution of the capacity matrix $\mathbf{H}\mathbf{H}^H$ ($C_n^2 = 5 \times 10^{-15} \text{ m}^{-2/3}$). This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

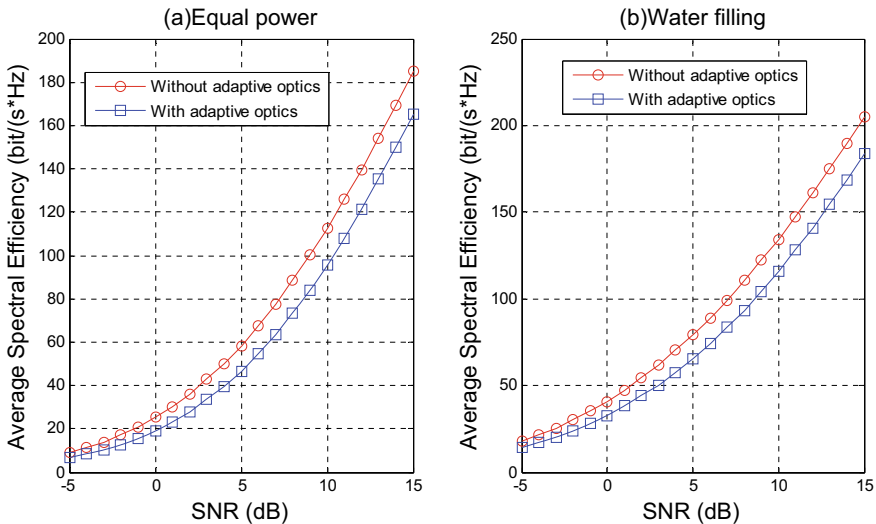


Fig. 5.40 The capacity of the 76 modes transmission system ($C_n^2 = 10^{-13} \text{ m}^{-2/3}$). **a** Without the channel information and equal power is assumed for each sub-channel. **b** With the channel information and the watering filling algorithm is implemented to optimize the power allocation. This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

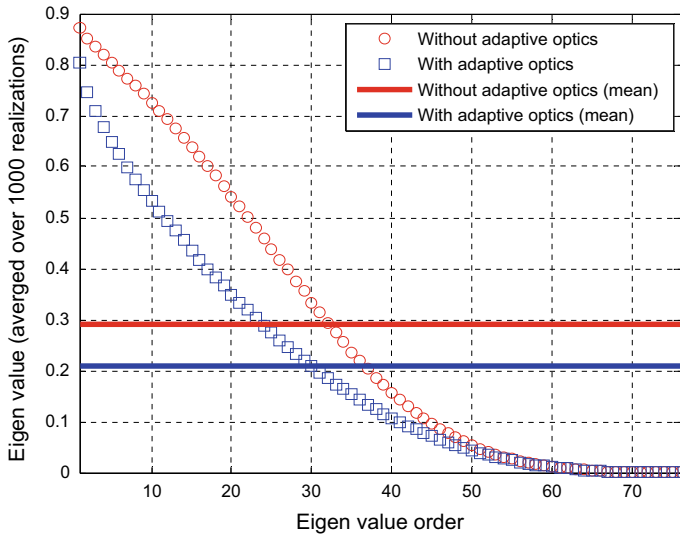


Fig. 5.41 The eigen value distribution of the capacity matrix $\mathbf{H}\mathbf{H}^H$ ($C_n^2 = 10^{-13} \text{ m}^{-2/3}$). This figure is reprinted/adapted with permission from [Opt. Express 26, 23008–23018 (2018)] © The Optical Society of America [95]

exacerbate the channel's impairments. Consequently, the use of adaptive optics in scenarios characterized by intense turbulence is not advisable.

5.6.4 Summary

In conclusion, our investigation into the capacity of free-space coherent MIMO systems, both with and without phase-only adaptive optics, has yielded insightful results. Under conditions of weak to moderate turbulence, the implementation of adaptive optics enhances system capacity in the absence of channel information. On the other hand, when comprehensive channel information is accessible, the relative performance of systems with and without adaptive optics depends on the SNR. Specifically, in the low SNR region, the non-adaptive-optics system exhibits superior performance, whereas in the high SNR domain, the adaptive optics enabled system surpasses its counterpart. However, in scenarios of intense turbulence, the application of phase-only adaptive optics consistently degrades system capacity, regardless of the availability of channel information. These findings offer valuable guidance for the future design and analysis of free-space coherent optical MIMO transmission systems, underscoring the importance of considering turbulence strength when deploying adaptive optics compensation strategies.

Appendix

Part of the section reprinted/adapted with permission from [Opt. Express 29, 13987–14005 (2021)] © The Optical Society of America [24].

According to [39], the elements of the matrix vector $\vec{\Lambda}$ are matrices generated using the rules described in Appendix A of [39]. It is re-summarized in the following using a different description language.

Denoting the three Pauli matrices as

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 0 & j \\ -j & 0 \end{pmatrix}, \quad (5.168)$$

There are three types of matrices to form the elements of the matrix vector $\vec{\Lambda}$:

1, $4N^2 - 4N$ matrices Λ_{3N+1} to Λ_{4N^2-N} are generated by setting the elements which do not belong to the 2×2 block matrix in the main diagonal, with 1 and 1 or j and $-j$ in the symmetrical or anti-symmetrical manner. The rest of the elements are set to be zero, with the normalization factor as $\sqrt{N}$.

2, $3N$ matrices Λ_1 to Λ_{3N} are generated by setting the three Pauli matrices in the main diagonal from the top left to the bottom right, with the normalization factor as $\sqrt{N}$.

Hence, $2N$ matrices related to the two Pauli matrices σ_2 and σ_3 can be attributed to the class 1 matrices which has the off-diagonal elements to be arranged with 1 and 1 or j and $-j$. Together, one has $4N^2 - 2N$ matrices.

N matrices are:

$$\sqrt{N} \begin{pmatrix} \ddots & & & & \\ & 0 & & & \\ & & 1 & & \\ & & & -1 & \\ & & & & 0 \\ & & & & & \ddots \end{pmatrix}, \quad (5.169)$$

3, $N - 1$ matrices are in the following form:

$$\Lambda_{4N^2-N+n} = \sqrt{\frac{N}{n(n+1)}} \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & -n & & \\ & & & & -n & \\ & & & & & 0 \\ & & & & & & 0 \\ & & & & & & & \ddots \end{pmatrix}, \quad (5.170)$$

$$n = 1 \cdots N - 1$$

4, Combining the matrix vector $\vec{\Lambda}$ with the last matrix Λ_0 , we have $\vec{\Lambda}'$ with $4N^2$ elements. Λ_0 is the identity matrix and it is with the normalization factor of 1.

Now considering the matrix

$$\mathbf{T} = \frac{1}{2N} \vec{\beta}' \cdot \vec{\Lambda}' = \frac{1}{2N} \sum_{n=0}^{4N^2-1} \beta_n \Lambda_n, \quad (5.171)$$

It is evident that the $4N^2-2N$ off-diagonal elements of the Hermitian symmetrical matrix $\mathbf{T}$ are generated by the $4N^2-4N$ matrices created following rule 1, and the $2N$ matrices created following rule 2 with the Pauli matrices σ_2 and σ_3 . Therefore, each off-diagonal element on the upper triangle of the matrix is formed by two independent and identically distributed Gaussian random processes, which are multiplied by 1 and j , i.e. $\frac{1}{2N} (\sqrt{N}\beta_l + j\sqrt{N}\beta_k)$. The lower triangle of the matrix is Hermitian symmetrical with respect to the upper triangle. Taking into account that $\langle \beta'_m(z) \beta'_n(z') \rangle = \frac{\sigma^2}{4N^2} \delta_{mn} \delta(z - z')$, we have:

$$\langle t_{mn}(z) t_{m'n'}^*(z') \rangle = \delta_{mm'} \delta_{nn'} \frac{\sigma^2}{8N^3} \delta(z - z'), \quad (5.172)$$

$$m \neq n, m' \neq n'$$

The off-diagonal elements are not correlated with the diagonal elements, suggesting

$$\langle t_{mn}(z) t_{m'm'}^*(z') \rangle = 0, \quad (5.173)$$

$$m \neq n$$

The diagonal terms can be computed by

$$\frac{1}{2N} \left(\sum_{n=0}^N \beta_n \Lambda_n + \sum_{n=4N^2-N+1}^{4N^2-1} \beta_n \Lambda_n \right), \quad (5.174)$$

Considering the $(2n + 1)$ th and the $(2n + 2)$ th diagonal elements, we have:

$$\begin{aligned}
 & t_{(2n+1)(2n+1)}(z) \\
 &= \frac{1}{2N} \left(\beta_0 + \sqrt{N} \beta_n + \sum_{k=n+1}^{N-1} \sqrt{\frac{N}{k(k+1)}} \beta_{4N^2-N+k} + \sqrt{\frac{N}{n(n+1)}} n \beta_{4N^2-N+n} \right) \\
 & t_{(2n+2)(2n+2)}(z) \\
 &= \frac{1}{2N} \left(\beta_0 - \sqrt{N} \beta_n + \sum_{k=n+1}^{N-1} \sqrt{\frac{N}{k(k+1)}} \beta_{4N^2-N+k} + \sqrt{\frac{N}{n(n+1)}} n \beta_{4N^2-N+n} \right),
 \end{aligned} \tag{5.175}$$

and therefore:

$$\begin{aligned}
 & \langle t_{(2n+1)(2n+1)}(z) t_{(2n+1)(2n+1)}^*(z') \rangle = \langle t_{(2n+2)(2n+2)}(z) t_{(2n+2)(2n+2)}^*(z') \rangle \\
 &= \frac{\sigma^2}{4N^2} \frac{1}{4N^2} \left(1 + N + \frac{nN}{(n+1)} + \sum_{k=n+1}^{N-1} \frac{N}{k(k+1)} \right) \delta(z - z') \\
 &= \frac{\sigma^2}{4N^2} \frac{1}{4N^2} 2N \delta(z - z') \\
 &= \frac{\sigma^2}{8N^3} \delta(z - z') \\
 & \langle t_{(2n+1)(2n+1)}(z) t_{(2n+2)(2n+2)}^*(z') \rangle \\
 &= \frac{\sigma^2}{4N^2} \frac{1}{4N^2} \left(1 - N + \frac{nN}{(n+1)} + \sum_{k=n+1}^{N-1} \frac{N}{k(k+1)} \right) \delta(z - z') \\
 &= 0,
 \end{aligned} \tag{5.176}$$

It is also easy to infer that the correlation function between the two non-adjacent main diagonal elements is zero.

Hence, one has:

$$\langle t_{mn}(z) t_{m'n'}^*(z') \rangle = \delta_{mn'} \delta_{nn'} \frac{\sigma^2}{8N^3} \delta(z - z'), \tag{5.177}$$

which is Eq. (5.57a). Therefore, we have:

$$\langle \mathbf{T}(z) \mathbf{T}^H(z') \rangle = \frac{\sigma^2}{4N^2} \mathbf{I} \delta(z - z'), \tag{5.178}$$

Assuming that we multiply matrix $\mathbf{T}$ with a small step Δz , by Eq. (5.178) and the property of Brownian motion, we have:

$$\begin{aligned}
\langle \mathbf{T} \Delta z \mathbf{T}^H \Delta z \rangle &\approx \left\langle \int_0^{\Delta z} \mathbf{T}(z) dz \int_0^{\Delta z} \mathbf{T}^H(z') dz' \right\rangle \\
&= \int_0^{\Delta z} dz \int_0^{\Delta z} \langle \mathbf{T}(z) \mathbf{T}^H(z') \rangle dz' \\
&= \frac{\sigma^2}{4N^2} \mathbf{I} \int_0^{\Delta z} dz \int_0^{\Delta z} \delta(z - z') dz' \\
&= \frac{\sigma^2}{4N^2} \mathbf{I} \Delta z,
\end{aligned} \tag{5.179}$$

Therefore, we have:

$$\langle \mathbf{T} dz \mathbf{T} dz \rangle = \langle \mathbf{T} dz \mathbf{T}^H dz \rangle = \frac{\sigma^2}{4N^2} \mathbf{I} dz, \tag{5.180}$$

which is Eq. (5.57b).

References

1. D. Donato et al., Stationary mode distribution and sidewall roughness effects in overmoded optical waveguides. *J. Lightwave Technol.* **28**(10), 1510–1520 (2010)
2. T. Barwicz, H.A. Haus, Three-dimensional analysis of scattering losses due to sidewall roughness in microphotonic waveguides. *J. Lightwave Technol.* **23**(9), 2719–2732 (2005)
3. J.H. Zhou, P. Gallion, Comprehensive analytical model to characterize randomness in optical waveguides. *Opt. Express* **24**(7), 6825–6842 (2016)
4. A. Snyder, J. Love, *Optical Waveguide Theory* (Springer, 1983)
5. D. Marcuse, Mode conversion caused by surface imperfections of a dielectric slab waveguide. *Bell Syst. Tech. J.* **48**(10), 3187–3215 (1969)
6. D. Marcuse, Derivation of coupled power equations. *Bell Syst. Tech. J.* **51**(1), 229–237 (1972)
7. D. Marcuse, Power distribution and radiation losses in multimode dielectric slab waveguides. *Bell Syst. Tech. J.* **51**(2), 429–454 (1972)
8. F. Ladouceur, J.D. Love, T.J. Senden, Effect of side wall roughness in buried channel waveguides. *IEE Proc. Optoelectron.* **141**(4), 242–248 (1994)
9. J.P.R. Lacey, F.P. Payne, Radiation loss from planar waveguides with random wall imperfections. *IEE Proc. J. (Optoelectron.)* **137**(4), 282–289 (1990)
10. D. Lenz, D. Erni, W. Bächtold, Modal power loss coefficients for highly overmoded rectangular dielectric waveguides based on free space modes. *Opt. Express* **12**(6), 1150–1156 (2004)
11. F. Grillot, L. Vivien, S. Laval, D. Pascal, E. Cassan, Size influence on the propagation loss induced by sidewall roughness in ultrasmall SOI waveguides. *IEEE Photonics Technol. Lett.* **16**(7), 1661–1663 (2004)
12. A.D. Simard, N. Ayotte, Y. Painchaud, S. Bedard, S. LaRochelle, Impact of sidewall roughness on integrated Bragg gratings. *J. Lightwave Technol.* **29**(24), 3693–3704 (2011)
13. D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic Press, New York, 1974)

14. K.P. Ho, J.M. Kahn, Mode coupling and its impact on spatially multiplexed systems. *Opt. Fiber Telecommun.* 491–568 (2013)
15. K.P. Ho, J.M. Kahn, Frequency diversity in mode-division multiplexing systems. *J. Lightwave Technol.* **29**(24), 3719–3726 (2011)
16. S.Ö. Arik, D. Askarov, J.M. Kahn, Effect of mode coupling on signal processing complexity in mode-division multiplexing. *J. Lightwave Technol.* **31**(3), 423–431 (2013)
17. P.J. Winzer, G.J. Foschini, MIMO capacities and outage probabilities in spatially multiplexed optical transport systems. *Opt. Express* **19**(17), 16680–16696 (2011)
18. M. Huang, X. Yan, Thermal-stress effects on the temperature sensitivity of optical waveguides. *J. Opt. Soc. Am. B* **20**(6), 1326–1333 (2003)
19. J.A. Anguita, M.A. Neifeld, B.V. Vasic, Spatial correlation and irradiance statistics in a multiple-beam terrestrial free-space optical communication link. *Appl. Opt.* **46**(26), 6561–6571 (2007)
20. P.K.A. Wai, C.R. Menyuk, Polarization mode dispersion, decorrelation, and diffusion in optical fibers with randomly varying birefringence. *J. Lightwave Technol.* **14**(2), 148–157 (1996)
21. A. Galtarossa, L. Palmieri, A. Pizzinat, B.S. Marks, C.R. Menyuk, An analytical formula for the mean differential group delay of randomly birefringent spun fibers. *J. Lightwave Technol.* **21**(7), 1635–1643 (2003)
22. Mecozzi, A theory of polarization-mode dispersion of spun fibers. *J. Lightwave Technol.* **27**(7), 938–943 (2009)
23. Q. Lin, G.P. Agrawal, Vector theory of stimulated Raman scattering and its application to fiber-based Raman amplifiers. *J. Opt. Soc. Am. B* **20**(8), 1616–1631 (2003)
24. J.H. Zhou, H.Q. Pu, Analytical study on the evolutionary behavior of transfer matrix element moments in strongly coupled multimode systems. *Opt. Express* **29**(9), 13987–14005 (2021)
25. D.J. Richardson, J.M. Fini, L.E. Nelson, Space-division multiplexing in optical fibres. *Nat. Photonics* **7**(5), 354–362 (2013)
26. Z. He et al., Independent component analysis based channel equalization for 6×6 MIMO-OFDM transmission over few-mode fiber. *Opt. Express* **24**(9), 9209–9217 (2016)
27. R. Ryf et al., Combined SDM and WDM transmission over 700-km few-mode fiber, in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)* (2013)
28. N. Bai, C. Xia, G. Li, Adaptive frequency-domain equalization for the transmission of the fundamental mode in a few-mode fiber. *Opt. Express* **20**(21), 24010–24017 (2012)
29. M.J. Murray, B. Redding, Quantitative strain sensing in a multimode fiber using dual frequency speckle pattern tracking. *Opt. Lett.* **45**(6), 1309–1312 (2020)
30. S. Buch, S. Mumtaz, R.-J. Essiambre, A.M. Tulino, G.P. Agrawal, Averaged nonlinear equations for multimode fibers valid in all regimes of random linear coupling. *Opt. Fiber Technol.* **48**, 123–127 (2019)
31. A. Mecozzi, C. Antonelli, M. Shtaif, Nonlinear propagation in multi-mode fibers in the strong coupling regime. *Opt. Express* **20**(11), 11673–11678 (2012)
32. Mecozzi, C. Antonelli, M. Shtaif, Coupled Manakov equations in multimode fibers with strongly coupled groups of modes. *Opt. Express* **20**(21), 23436–23441 (2012)
33. S. Mumtaz, R.J. Essiambre, G.P. Agrawal, Nonlinear propagation in multimode and multicore fibers: generalization of the Manakov equations. *J. Lightwave Technol.* **31**(3), 398–406 (2013)
34. Antonelli, M. Shtaif, A. Mecozzi, Modeling of nonlinear propagation in space-division multiplexed fiber-optic transmission. *J. Lightwave Technol.* **34**(1), 36–54 (2016)
35. F. Poletti, P. Horak, Description of ultrashort pulse propagation in multimode optical fibers. *J. Opt. Soc. Am. B* **25**(10), 1645–1654 (2008)
36. Antonelli, A. Mecozzi, M. Shtaif, Raman amplification in multimode fibers with random mode coupling. *Opt. Lett.* **38**(8), 1188–1190 (2013)
37. Z. Zhang, Y. Lu, Y. Pan, X. Bao, L. Chen, Trench-assisted multimode fiber used in Brillouin optical time domain sensors. *Opt. Express* **27**(8), 11396–11405 (2019)

38. A. Zadok, E. Zilka, A. Eyal, L. Thévenaz, M. Tur, Vector analysis of stimulated Brillouin scattering amplification in standard single-mode fibers. *Opt. Express* **16**(26), 21692–21707 (2008)
39. A. Antonelli, A. Mecozzi, M. Shtaif, P.J. Winzer, Stokes-space analysis of modal dispersion in fibers with multiple mode transmission. *Opt. Express* **20**(11), 11718–11733 (2012)
40. Q. Hu, W. Shieh, Autocorrelation function of channel matrix in few-mode fibers with strong mode coupling. *Opt. Express* **21**(19), 22153–22165 (2013)
41. A. Andrusier, M. Shtaif, C. Antonelli, A. Mecozzi, Assessing the effects of mode-dependent loss in space-division multiplexed systems. *J. Lightwave Technol.* **32**(7), 1317–1322 (2014)
42. K.P. Ho, J.M. Kahn, Statistics of group delays in multimode fiber with strong mode coupling. *J. Lightwave Technol.* **29**(21), 3119–3128 (2011)
43. P. Sillard, D. Molin, M. Bigot-Astruc, K. de Jongh, F. Achten, Low-differential-mode-group-delay 9-LP-mode fiber, in *Optical Fiber Communication Conference* (Optica Publishing Group, Los Angeles, California, 2015) in OSA Technical Digest (online)
44. S. Berdagué, P. Facq, Mode division multiplexing in optical fibers. *Appl. Opt.* **21**(11), 1950–1955 (1982)
45. M.B. Shemirani, W. Mao, R.A. Panicker, J.M. Kahn, Principal modes in graded-index multi-mode fiber in presence of spatial- and polarization-mode coupling. *J. Lightwave Technol.* **27**(10), 1248–1261 (2009)
46. K. Kitayama, S. Seikai, N. Uchida, Impulse response prediction based on experimental mode coupling coefficient in a 10-km long graded-index fiber. *IEEE J. Quantum Electron.* **16**(3), 356–362 (1980)
47. T. Antonia, V. Sergio, *Random Matrix Theory and Wireless Communications* (Commun. Inf. Theory) (Now Foundations and Trends, 2004)
48. J. Zhou, Q. Hu, Higher order statistics of the Mueller matrix in a fiber with an arbitrary length impacted by PMD. *Opt. Express* **28**(20), 30063–30074 (2020)
49. A. Mecozzi, C. Antonelli, M. Shtaif, Intensity impulse response of SDM links. *Opt. Express* **23**(5), 5738–5743 (2015)
50. A. Antonelli, A. Mecozzi, M. Shtaif, N.K. Fontaine, H. Chen, R. Ryf, Stokes-space analysis of modal dispersion of SDM fibers with mode-dependent loss: theory and experiments. *J. Lightwave Technol.* **38**(7), 1668–1677 (2020)
51. G.J. Foschini, C.D. Poole, Statistical theory of polarization dispersion in single mode fibers. *J. Lightwave Technol.* **9**(11), 1439–1456 (1991)
52. Y. Li, A. Yariv, Solutions to the dynamical equation of polarization-mode dispersion and polarization-dependent losses. *J. Opt. Soc. Am. B* **17**(11), 1821–1827 (2000)
53. A.J. Laub, *Matrix Analysis—For Scientists and Engineers* (SIAM, Chap. 13, 2004)
54. Klain, The Matrix Exponential and Linear Systems of ODEs (2019). [Online]. <http://faculty.uml.edu/dklain/exponential.pdf>
55. J.H. Zhou, H.Q. Pu, Analytical expressions for the crosstalk of super-modes in the tightly bounded multicore fibers. *Opt. Express* **30**(4), 4833–4844 (2022)
56. J.D. Downie, X. Liang, S. Makovejs, Assessing capacity and cost/capacity of 4-core multicore fibers against single core fibers in submarine cable systems. *J. Lightwave Technol.* **38**(11), 3015–3022 (2020)
57. G. Rademacher et al., High capacity transmission in a coupled-core three-core multi-core fiber. *J. Lightwave Technol.* **39**(3), 757–762 (2021)
58. Yahav, N. Sheffi, Y. Biofcic, D. Sadot, Multi-gigabit spatial-division multiplexing transmission over multicore plastic optical fiber. *J. Lightwave Technol.* **39**(8), 2296–2304 (2021)
59. Antonelli, A. Mecozzi, M. Shtaif, The delay spread in fibers for SDM transmission: dependence on fiber parameters and perturbations. *Opt. Express* **23**(3), 2196–2202 (2015)
60. Saitoh, S. Matsuo, Multicore fiber technology. *J. Lightwave Technol.* **34**(1), 55–66 (2016)
61. M. Fini, B. Zhu, T.F. Taunay, M.F. Yan, Statistics of crosstalk in bent multicore fibers. *Opt. Express* **18**(14), 15122–15129 (2010)
62. T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, E. Sasaoka, Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber. *Opt. Express* **19**(17), 16576–16592 (2011)

63. Koshiba, K. Saitoh, K. Takenaga, S. Matsuo, Multi-core fiber design and analysis: coupled-mode theory and coupled-power theory. *Opt. Express* **19**(26), B102–B111 (2011)
64. J.M. Fini, B. Zhu, T.F. Taunay, M.F. Yan, K.S. Abedin, Crosstalk in multicore fibers with randomness: gradual drift vs. short-length variations. *Opt. Express* **20**(2), 949–959 (2012)
65. T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, E. Sasaoka, Characterization of crosstalk in ultra-low-crosstalk multi-core fiber. *J. Lightwave Technol.* **30**(4), 583–589 (2012)
66. A. Macho, M. Morant, R. Llorente, Experimental evaluation of nonlinear crosstalk in multi-core fiber. *Opt. Express* **23**(14), 18712–18720 (2015)
67. M.-J. Li, S. Li, R.A. Modavis, Coupled mode analysis of crosstalk in multicore fiber with random perturbations, in *Optical Fiber Communication Conference* (Optica Publishing Group, Los Angeles, California, 2015) in OSA Technical Digest (online)
68. A.V.T. Cartaxo, T.M.F. Alves, Discrete changes model of inter-core crosstalk of real homogeneous multi-core fibers. *J. Lightwave Technol.* **35**(12), 2398–2408 (2017)
69. T.M.F. Alves, A.V.T. Cartaxo, Characterization of the stochastic time evolution of short-term average intercore crosstalk in multicore fibers with multiple interfering cores. *Opt. Express* **26**(4), 4605–4620 (2018)
70. A.V.T. Cartaxo, J.A.P. Morgado, New expression for evaluating the mean crosstalk power in weakly-coupled multi-core fibers. *J. Lightwave Technol.* **39**(6), 1830–1842 (2021)
71. A. Antonelli, G. Riccardi, T. Hayashi, A. Mecozzi, Role of polarization-mode coupling in the crosstalk between cores of weakly-coupled multi-core fibers. *Opt. Express* **28**(9), 12847–12861 (2020)
72. S. Saitoh, K. Takenaga, K. Aikawa, Demonstration of a rectangularly-arranged strongly-coupled multi-core fiber, in *Optical Fiber Communication Conference* (Optica Publishing Group, San Diego, California, 2018) in OSA Technical Digest (online)
73. Xia, N. Bai, I. Ozdur, X. Zhou, G. Li, Supermodes for optical transmission. *Opt. Express* **19**(17), 16653–16664 (2011)
74. X. Cen et al., Supermodes in strongly-coupled multi-core fibers, in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, 17–21 March 2013 (2013), pp. 1–3
75. S. Li, J. Wang, Supermode fiber for orbital angular momentum (OAM) transmission. *Opt. Express* **23**(14), 18736–18745 (2015)
76. Xia et al., Supermodes in coupled multi-core waveguide structures. *IEEE J. Sel. Top. Quantum Electron.* **22**(2), 196–207 (2016)
77. J. Zhou, H. Pu, Generalized analytical study on the random crosstalk in multicore/multimode fibers. *J. Lightwave Technol.* **40**(4), 1112–1120 (2022)
78. T. Sakamoto, T. Mori, M. Wada, T. Yamamoto, K. Nakajima, Coupled single-mode multi-core fiber design for long-haul MIMO transmission system, in *Optical Fiber Communication Conference* (Optica Publishing Group, Los Angeles, California, 2017) in OSA Technical Digest (online)
79. Ip et al., SDM transmission of real-time 10GbE traffic using commercial SFP + transceivers over 0.5km elliptical-core few-mode fiber. *Opt. Express* **23**(13), 17120–17126 (2015)
80. S. Chen, J. Wang, PANDA-type elliptical-core multi-mode fiber with fully lifted Eigenmodes for low-crosstalk direct fiber vector Eigenmode space-division multiplexing, in *Optical Fiber Communication Conference* (Optica Publishing Group, San Diego, California, 2018) in OSA Technical Digest (online)
81. S. Fu et al., Panda type few-mode fiber capable of both mode profile and polarization maintenance. *J. Lightwave Technol.* **36**(24), 5780–5785 (2018)
82. J.H. Zhou, J.B. Zong, D.Q. Liu, Coupled mode theory for orbital angular momentum modes transmission in the presence of atmosphere turbulence. *Opt. Express* **23**(25), 31964–31976 (2015)
83. J. Wang et al., Terabit free-space data transmission employing orbital angular momentum multiplexing. *Nat. Photonics* **6**(7), 488–496 (2012)

84. H. Huang et al., 100 Tbit/s free-space data link using orbital angular momentum mode division multiplexing combined with wavelength division multiplexing, in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013* (Optica Publishing Group, Anaheim, California, 2013) in OSA Technical Digest (online)
85. Guan et al., Free-space coherent optical communication with orbital angular, momentum multiplexing/demultiplexing using a hybrid 3D photonic integrated circuit. *Opt. Express* **22**(1), 145–156 (2014)
86. Bozinovic, P. Kristensen, S. Ramachandran, Long-range fiber-transmission of photons with orbital angular momentum, in *CLEO: 2011—Laser Science to Photonic Applications* (2011)
87. Paterson, Atmospheric turbulence and orbital angular momentum of single photons for optical communication. *Phys. Rev. Lett.* **94**(15), 153901 (2005)
88. J.A. Anguita, M.A. Neifeld, B.V. Vasic, Turbulence-induced channel crosstalk in an orbital angular momentum-multiplexed free-space optical link. *Appl. Opt.* **47**(13), 2414–2429 (2008)
89. A. Tyler, R.W. Boyd, Influence of atmospheric turbulence on the propagation of quantum states of light carrying orbital angular momentum. *Opt. Lett.* **34**(2), 142–144 (2009)
90. B. Rodenburg et al., Influence of atmospheric turbulence on states of light carrying orbital angular momentum. *Opt. Lett.* **37**(17), 3735–3737 (2012)
91. Y. Zhang et al., Influence of atmospheric turbulence on the transmission of orbital angular momentum for Whittaker-Gaussian laser beams. *Opt. Express* **22**(18), 22101–22110 (2014)
92. A. Belmonte, Feasibility study for the simulation of beam propagation: consideration of coherent lidar performance. *Appl. Opt.* **39**(30), 5426–5445 (2000)
93. V. Oppenheim, A.S. Willsky, S.H. Nawab, *Signals and Systems*, 2nd edn. (Prentice Hall, 1997)
94. V. P. Aksenov, V. V. Kolosov, Scintillation index of a laser beam having the initial orbital angular momentum in the turbulent atmosphere, in *Imaging and Applied Optics 2015* (Optica Publishing Group, Arlington, Virginia, 2015) in OSA Technical Digest (online)
95. J.H. Zhou, J.J. Wu, Q.S. Hu, Capacity analysis for free space coherent optical MIMO transmission systems: with and without adaptive optics. *Opt. Express* **26**(18), 23008–23018 (2018)
96. Zhao, X. Li, G. Li, J.M. Kahn, Capacity limits of spatially multiplexed free-space communication. *Nat. Photonics* **9**(12), 822–826 (2015)
97. K. Zhu, G. Zhou, X. Li, X. Zheng, H. Tang, Propagation of Bessel-Gaussian beams with optical vortices in turbulent atmosphere. *Opt. Express* **16**(26), 21315–21320 (2008)
98. J. Zhou, J. Zong, D. Liu, The higher order statistics of OAM modal amplitudes under atmosphere turbulence. *IEEE Photonics Technol. Lett.* **28**(23), 2653–2656 (2016)
99. Chandrasekaran, J.H. Shapiro, Photon information efficient communication through atmospheric turbulence—part I: channel model and propagation statistics. *J. Lightwave Technol.* **32**(6), 1075–1087 (2014)
100. N. Chandrasekaran, J.H. Shapiro, L. Wang, Photon information efficient communication through atmospheric turbulence—part II: bounds on ergodic classical and private capacities. *J. Lightwave Technol.* **32**(6), 1088–1097 (2014)
101. J. Zhou, J. Wu, J. Zong, Q. Hu, Optimal mode set selection for free space optical communications in the presence of atmosphere turbulence. *J. Lightwave Technol.* **36**(11), 2222–2229 (2018)
102. Y. Ren et al., Adaptive-optics-based simultaneous pre- and post-turbulence compensation of multiple orbital-angular-momentum beams in a bidirectional free-space optical link. *Optica* **1**(6), 376–382 (2014)
103. M. Li, M. Cvijetic, Y. Takashima, Z. Yu, Evaluation of channel capacities of OAM-based FSO link with real-time wavefront correction by adaptive optics. *Opt. Express* **22**(25), 31337–31346 (2014)
104. Y. Ren et al., Turbulence compensation of an orbital angular momentum and polarization-multiplexed link using a data-carrying beacon on a separate wavelength. *Opt. Lett.* **40**(10), 2249–2252 (2015)
105. A. Lapidoth, S.M. Moser, M.A. Wigger, On the capacity of free-space optical intensity channels. *IEEE Trans. Inf. Theory* **55**(10), 4449–4461 (2009)

106. J.B. Wang, Q.S. Hu, J. Wang, M. Chen, J.Y. Wang, Tight bounds on channel capacity for dimmable visible light communications. *J. Lightwave Technol.* **31**(23), 3771–3779 (2013)
107. J.-Y. Wang, J. Dai, R. Guan, L. Jia, Y. Wang, M. Chen, Channel capacity and receiver deployment optimization for multi-input multi-output visible light communications. *Opt. Express* **24**(12), 13060–13074 (2016)
108. S.M. Moser, Capacity results of an optical intensity channel with input-dependent Gaussian noise. *IEEE Trans. Inf. Theory* **58**(1), 207–223 (2012)
109. Tse, P. Viswanath, *Fundamentals of Wireless Communication* (Cambridge University Press, Cambridge, 2005)

Chapter 6

Mode Generation and Multiplexing for Multi-mode Waveguides and Free Space



Abstract Mode generation and multiplexing for multi-mode waveguides and free space. The chapter studies the mode generator and multiplexer, which are the key components for mode division multiplexing systems. The chapter discusses the novel ways to realize mode generators and multiplexers.

Mode generator and multiplexer are the key components for mode division multiplexing systems. The chapter discusses novel ways to realize mode generators and multiplexers. Multiplexers for multi-core fiber supermodes, ring core fibers and multimode cylindrical fiber are proposed, which are based on the physical propagation characteristics. Finally, a multiplexer for the spatial orbital angular momentum (OAM) modes is proposed based on the concept of angular lens.

6.1 A Novel Mode Multiplexer/De-Multiplexer for Multi-core Fibers

Part of the section reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1].

In this section, we introduce a mode multiplexer/de-multiplexer (MUX/DEMUX) design for multi-core fibers (MCFs) based on the phenomenon of multimode interference (MMI). This innovative approach exploits the close match between the MMI coupler's transfer matrix rows with the linearly arranged core modal profiles of MCFs. By incorporating a phase shifter array, the MMI coupler effectively discriminates between various super-modes present in the MCFs, with the output port number directly corresponding to the super-mode index. Comprehensive theoretical analyses and numerical simulations are presented to demonstrate the multiplexing and de-multiplexing capabilities of this device. Notably, the proposed mode MUX/DEMUX can be realized either in fiber or waveguide form, facilitating the integration with the fiber-optic systems.

6.1.1 Introduction

Recently, Mode Division Multiple Access (MDMA) has gained considerable interest owing to its potential for significantly enhancing transmission capacity [2–4]. Theoretically, the multimode transmission system boosts the capacity so that it is multiplied by the number of transmission modes in comparison to the single-mode case. Among the strategies to realize such a system, the employment of Multi-Core Fibers (MCFs), which are characterized by their multiple cores designed for signal propagation, offer a promising opportunity for realizing multimode transmission [4–7].

Essential components for MCF implementation include mode MUXs/DEMUXs tailored specifically for MCFs. While free space optics can distinguish various modes [4], it is of considerable complexity and cost, particularly when identifying higher-order modes. Alternatively, the phase matching principle offers an efficient fiber-based approach for mode de-multiplexing [6]. However, designing such a device for MCFs with a substantial number of cores proves challenging.

The other potential method for designing MCF mode MUX/DEMUX is to implement the arrayed waveguide gratings (AWGs) [7]. The super-mode order within the MCF corresponds directly to the AWG's output port number. This groundbreaking proposal necessitates the use of planar lightwave circuit (PLC) technology for AWG fabrication, which makes integration with the fiber system difficult [6]. Moreover, the AWG's internal field interference requires equal wave amplitudes at the input ports, whereas the super-mode field distribution within the MCF follows the sinusoidal pattern [8]. Consequently, as the core count increases, the modal extinction ratio of AWG-based MUXs/DEMUXs may suffer degradation.

An alternative approach to manipulate modes is to exploit the Multimode Interference (MMI) effect [9–14], which has demonstrated capabilities in generating [13] and converting [14] modes within multimode waveguides. In this study, we introduce a novel method to implement mode MUXs/DEMUXs for MCFs utilizing MMI couplers constructed from fibers with rectangular cores or multimode waveguides based on PLC technology. The concept can be summarized as follows: the super-modes propagating within the MCF are directed into an MMI coupler. The amplitudes of the waves in the MCF cores obey the sinusoidal function, and their phases are adjusted using a phase shifter array. These phase shifters modify the transfer matrix of the multimode waveguide [10] aligning its rows with the amplitude vectors of the MCF's super-modes. When a specific super-mode is introduced, it undergoes phase shifting and is transmitted into the multimode waveguide. Consequently, the corresponding output port of the MMI coupler outputs the optical wave, while the rest ports remain inactive. This mechanism efficiently realizes the MUX/DEMUX functionality.

It is noteworthy that this approach enables the multiplexing and de-multiplexing of modes with distinct amplitude and phase distributions. However, it is acknowledged that polarization-induced mode diversity is not explicitly addressed in this section.

Nevertheless, the multiplexing and de-multiplexing of TM and TE modes have been extensively discussed in the existing literatures [15].

6.1.2 Mathematical Description

6.1.2.1 Super-Modes Inside the MCF

Within this work, we analyze under the assumption that the identical cores of the MCF are linearly arranged and uniformly spaced, which is the same as the previous studies in [6, 7]. Assuming that each individual core supports a single mode, a system of $N-1$ coupled cores will exhibit $N-1$ super-modes. Notably, the amplitudes of these super-modes on each core follows the sinusoidal function [8]:

$$\sin\left(\frac{np\pi}{N}\right) \quad (6.1)$$

where n is the number of the super-mode, p the number of the cores, both being integers ranging from 1 to $N-1$, where N is the total number of cores. The diameter of each core is assumed to be a , and the distance between the adjacent core centers is assumed to be D .

6.1.2.2 MCF Mode MUX-DEMUX Based on MMI Couplers

Assuming that the MCF cores are connected with a fiber possessing a rectangular core or a rectangular multimode waveguide, built by the PLC technology, multimode interference naturally arises. For the sake of generality, let's consider a rectangular core fiber serving as the MMI coupler, and the subsequent analysis will focus on the fiber technology-based mode MUX-DEMUX. This fiber possesses a height of a and an effective width of ND . The rectangular fiber is single-mode in the y -direction and multimode in the x -direction. Consequently, the modal contribution along the y -direction is represented by an effective index. Within the multimode waveguide, the x -dimensional modes can be mathematically described by sinusoidal functions [9–13], which align with Eq. (6.1). Furthermore, we assume that the MCF cores, interfaced with the MMI coupler, are extremely thin, approaching a limit where a tends to be 0. Consequently, the modal distribution function of these cores can be idealized as the Dirac function. It is noteworthy that this assumption can be extended to thicker cores, which will be discussed later in this section. Each core is added with a phase shifter, introducing a phase shift of $\frac{i^2\pi}{2N}$, where i corresponds to the port number. The center of the cores of the MCF are placed at the following positions [10]

$$x_i = \frac{iW_e}{N} = iD \quad (i = 1 \dots N - 1) \quad (6.2)$$

where W_e stands for the effective width of the MMI coupler. We assume that the n th super-mode of MCF is injected at the input and the mode experience phase shifted, the x dimensional modal field should be expressed as:

$$E(0, x) = \sum_{i=1}^{N-1} \sin\left(\frac{ni\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \delta(x - x_i) \quad (6.3)$$

We set the MMI coupler length as [12]

$$L_N = \frac{2n_{eff} W_e^2}{N\lambda_0} \quad (6.4)$$

where n_{eff} stands for the effective index of the fundamental mode of the MMI coupler and λ_0 represent the free space wavelength. The x dimensional mode propagation inside the multimode waveguide can be formulated as [10]

$$E(z, x) = \exp(-jkz) \sum_{m=1}^{\infty} E_m \exp\left(j\frac{m^2\pi\lambda_0 z}{4n_{eff} W_e^2}\right) \sin\left(\frac{m\pi x}{W_e}\right) \quad (6.5)$$

where E_m stands for the m^{th} mode amplitude. In case the input field follows Eq. (6.3), one may express E_m as

$$E_m = \frac{2}{W_e} \sum_{i=0}^{N-1} \sin\left(\frac{ni\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \sin\left(\frac{mi\pi}{N}\right) \quad (6.6)$$

At the propagation distance L_N , the field becomes

$$\begin{aligned} E(L_N, x) &= \frac{2}{W_e} \exp(-jkL_N) \\ &\times \sum_{m=1}^{\infty} \sin\left(\frac{m\pi x}{W_e}\right) \exp\left(j\frac{m^2\pi}{2N}\right) \sum_{i=1}^{N-1} \sin\left(\frac{ni\pi}{N}\right) \sin\left(\frac{mi\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \end{aligned} \quad (6.7)$$

Let $m = lN + n_0$, where l stands for an integer and n_0 stands for an integer between 0 and $N-1$, and consider the last three terms in Eq. (6.7), we have:

$$\begin{aligned} &\sum_{i=1}^{N-1} \sin\left(\frac{mi\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \sin\left(\frac{ni\pi}{N}\right) \\ &= \sum_{i=1}^{N-1} \sin\left(\frac{n_0 i\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \sin\left(\frac{ni\pi}{N}\right) \end{aligned}$$

$$= \sqrt{\frac{N}{2}} j \exp\left(j\frac{\pi}{4}\right) \sin\left(\frac{n_0 n \pi}{N}\right) \exp\left(-j\frac{(n_0^2 + n^2)\pi}{2N}\right) \quad (6.8)$$

Combining Eqs. (6.7) and (6.8) and using $m = lN + n_0$, we have:

$$\begin{aligned} E(L_N, x) &= \exp(-jkL_N) \sqrt{\frac{N}{2}} j \exp\left(j\frac{\pi}{4}\right) \exp\left(-j\frac{n^2\pi}{2N}\right) \\ &\times \frac{2}{W_e} \sum_{m=1}^{\infty} \sin\left(\frac{m\pi x}{W_e}\right) \sin\left(\frac{mn\pi}{N}\right) \end{aligned} \quad (6.9)$$

The summation term in Eq. (6.9) can be simplified as:

$$\begin{aligned} \frac{2}{W_e} \sum_{m=1}^{\infty} \sin\left(\frac{mn\pi}{N}\right) \sin\left(\frac{m\pi x}{W_e}\right) &= \delta(x - x_n) \\ x_n &= n \frac{W_e}{N} \end{aligned} \quad (6.10)$$

Therefore, we have Eq. (6.10) to be converted as:

$$\begin{aligned} E(L_N, x) &= \exp(-jkL_N) j \exp\left(j\frac{\pi}{4}\right) \\ &\times \sqrt{\frac{N}{2}} \exp\left(-j\frac{n^2\pi}{2N}\right) \delta(x - x_n) \end{aligned} \quad (6.11)$$

It can be inferred from Eq. (6.11) that the output field occurs at the n^{th} output port of the MMI coupler with the phase shift of $-j\frac{n^2\pi}{2N}$ if one ignores the constant phase term. Hence, mode n at the input has resulted the optical intensity at the n^{th} output port. It is notable that since the mode profile is within the range of $[0, W_e]$, n should be an integer ranging from 1 to $N-1$, which is in agreement with the modal number within the MCF.

Now, let's consider MCFs that possess thick cores. Presuming the core operates in a monomode regime and its field profile along the x-dimension is denoted by $f(x)$, we will consider the input as:

$$E(0, x) = \left(\sum_{i=1}^{N-1} \sin\left(\frac{ni\pi}{N}\right) \exp\left(j\frac{i^2\pi}{2N}\right) \delta(x - x_i) \right) \otimes f(x) \quad (6.12)$$

It can be seen from the equation that the first three terms in Eq. (6.12) are in exact agreement with that in Eq. (6.3). Due to the fact that the MMI coupler is a linear system, the output field can be expressed as:

$$E(L_N, x) = \exp(-jkL_N) j \exp\left(j\frac{\pi}{4}\right)$$

$$\times \sqrt{\frac{N}{2}} \exp\left(-j \frac{n^2 \pi}{2N}\right) f(x - x_n) \quad (6.13)$$

Therefore, it can be inferred from Eq. (6.13) that the output field occurs on the n th output port of if the actual size of the core is taken into account.

6.1.3 Numerical Results and Discussion

Numerical simulations utilizing the three-dimensional beam propagation method (3DBPM) have been conducted for the envisioned device. To illustrate, a seven-core multi-core fiber (MCF) is employed as a test case in these simulations. This MCF has an index profile and geometric characteristics like those reported in References [6, 7], which are summarized here for the convenience of readers. Specifically, the cladding index is set at 1.45, while the index difference between the cores and cladding is 1.2%. Each core possesses a diameter of 5 μm , and the seven cores are uniformly spaced and linearly arranged. To evaluate the robustness of the methodology, three distinct inter-core center-to-center distances (D) are considered: 5, 7, and 10 μm .

As is mentioned in the preceding section, the rectangular multimode fiber employed in the study possesses a height of 5 μm and an effective width that scales with the inter-core distance D , specifically $8D$, resulting in widths of 40, 56, and 80 μm for the respective D values of 5, 7, and 10 μm . This rectangular multimode fiber shares the same core and cladding refractive indices as the MCF. The initial length of the MMI coupler is approximately estimated using Eq. (6.4) and adjusted during the calculation to get the best performance. A schematic diagram of the device under consideration is depicted in Fig. 6.1.

There exist multiple approaches to realize phase shifters. These include incorporating an external phase shifter array [16], utilizing an array of bent waveguides to introduce phase shifts between input and output planes [7], or fabricating a MCF where each core possesses a unique refractive index [6]. Notably, the first two

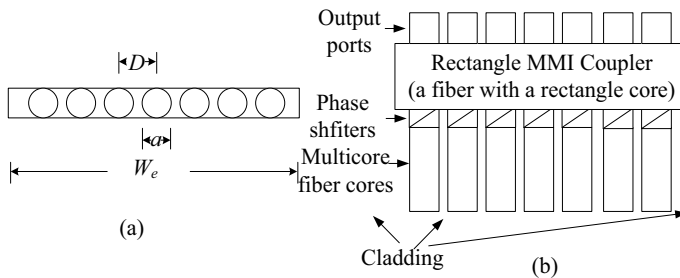


Fig. 6.1 Schematic of the device. **a** Cross section of the MCF and MMI coupler. **b** Top view of the device. This figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1]

methods inherently prevent coupling between fields within distinct cores. Similarly, even with the third method, coupling can be effectively neglected since the waveguides can be spatially separated during phase shift introduction. For the sake of generality, we assume the utilization of a perfect external phase shifter array that precisely achieves the desired phase shift in each core, without any cross-talk among them. More specifically speaking, the fields within the cores are assumed to propagate with specific lengths without coupling from other cores in the simulation.

Firstly, we validate and assess the robustness of the proposed device by varying the core center distance (D) and simulating the phase-shifted fundamental modes of different multi-core fibers (MCFs) being launched into the MMI coupler. Figure 6.2 depicts the field profiles of the fundamental super-mode and the output fields from the MMI coupler for various D values. As is evident from the figure, the modal profiles exhibit sinusoidal patterns, as is anticipated in [8], suggesting that an arrayed waveguide grating (AWG) might not be ideally suited for mode multiplexing/demultiplexing (MUX/DEMUX). Also, it is illustrated that the output fields predominantly concentrate at the first output port of the MMI couplers, showing the device's potential as an effective mode MUX/DEMUX. The modal extinction ratio, defined as the power ratio between the first port and all other ports, exceeds 30 dB, indicating excellent performance. Comparative analysis reveals no significant performance differences among the three scenarios with varying D values.

Given the above analysis, it is reasonable to select the case with $D = 7 \mu\text{m}$ as a representative example to show the device's performance for higher-order super-modes. Figures 6.3 and 6.4 illustrate the field profiles of the 2nd to 7th super-modes within the MCF, along with the corresponding output fields from the MMI couplers. As the mode number increases, the output fields gradually shift from left to right, aligning with our predictions. Notably, the modal extinction ratio remains above 30 dB for all modes, further confirming the device's robust performance across different super-modes.

To underscore the device's mode demultiplexing capability, we consider a scenario where a mixed field comprising the 1st and 2nd modes propagates through the multi-core fiber (MCF). Specifically, the 1st mode is assigned an amplitude of 2, whereas the 2nd mode has an amplitude of 1, with both modes sharing the same phase. The combined field is illustrated in Fig. 6.5a. The field is phase shifted and passed on to the MMI coupler. The mixed field is effectively separated, with the 1st and 2nd modes emerging at their respective output ports with accurate amplitudes, as is depicted in Fig. 6.5b. This example shows the device's proficiency in demultiplexing different modes.

The preceding discussion underscores the device's capability for mode demultiplexing. To achieve multiplexing, a straightforward approach involves reversing the roles of the output and input ports. By utilizing the output ports as input ports and the output fields as the input fields, it can be readily inferred that this configuration enables the generation of the super-modes of the MCF.

The fabrication error tolerance of the device is a crucial aspect that has been thoroughly investigated. Assuming the same parameters as in the case of $D = 7 \mu\text{m}$, two

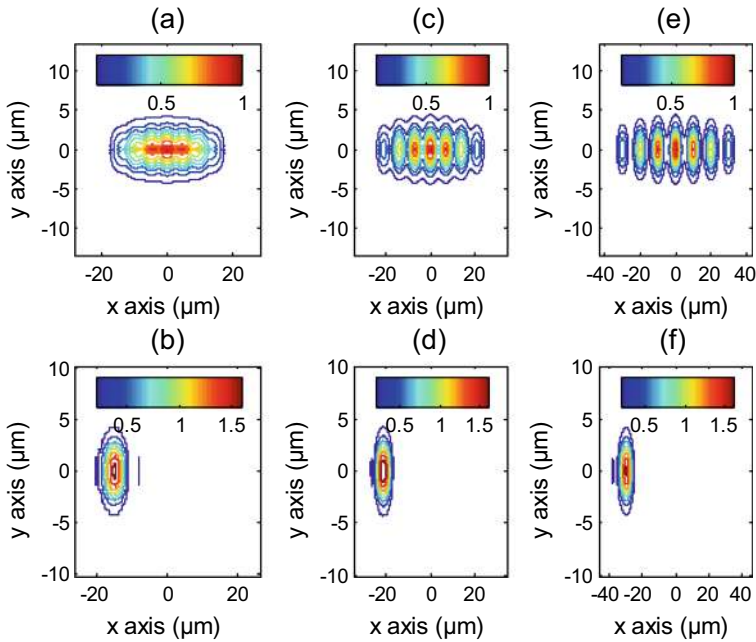


Fig. 6.2 The fundamental super-mode of the MCF and the field at the output of the DEMUX with different core distances, **a** the fundamental mode when $D = 5 \mu\text{m}$, **b** the field at the output of the DEMUX when $D = 5 \mu\text{m}$, **c** the fundamental mode when $D = 7 \mu\text{m}$, **d** the field at the output of the DEMUX when $D = 7 \mu\text{m}$, **e** the fundamental mode when $D = 10 \mu\text{m}$, **f** the field at the output of the DEMUX when $D = 10 \mu\text{m}$. This figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1]

primary error types were analyzed: phase errors in the phase shifter array and fabrication errors in the MMI coupler. A 10% phase error in the phase shifter array, where each phase term increases by 10%, leads to a modal extinction ratio degradation to 12 dB. Meanwhile, a $40 \mu\text{m}$ length error (approximately 5% of the total length) or a $1.4 \mu\text{m}$ width error (around 2.5% of the total width) in the MMI coupler results in a more significant degradation of the modal extinction ratio to 25 dB and 26 dB, respectively. As is indicated in [11], errors in the width and length of the MMI coupler are proportionally related, with width errors having a more pronounced impact on performance compared to length errors [11]. To mitigate these impacts, precise fabrication of the phase shifter array is essential, or the design should incorporate adjustability through electro-optical tuning during application. Additionally, careful adjustment of the MMI coupler's length, similar to the approach taken in multimode fiber-based bandpass filters, is crucial to ensure optimal performance [17].

Finally, temperature variations and stress can influence the refractive index of the device. However, for fiber-based devices constructed from silica, these effects are negligible, as is verified by BPM simulations.

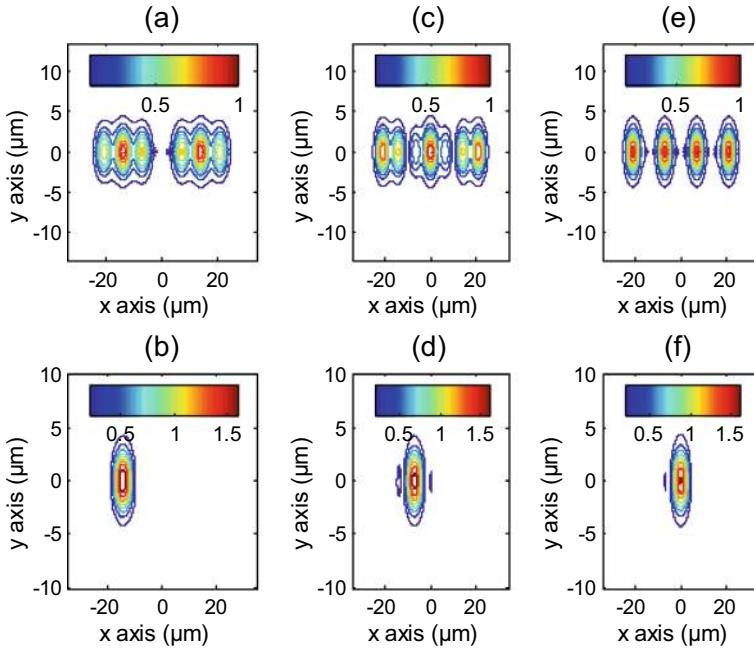


Fig. 6.3 The super-modes of the MCF and the fields at the output of the DEMUX with $D = 7 \mu\text{m}$, **a** the 2nd mode, **b** the field at the output of the DEMUX for the 2nd mode, **c** the 3rd mode, **d** the field at the output of the DEMUX for the 3rd mode, **e** the 4th mode, **f** the field at the output of the DEMUX for the 4th mode. This figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1]

6.1.4 Summary

In conclusion, a novel method based on the multimode interference effect is introduced to realize mode MUX/DEMUX in MCFs. This device can be implemented through an either fiber-based or planar lightwave circuit (PLC) design, offering compactness and compatibility with MCFs with an arbitrary number of linearly aligned cores. Additionally, the proposed device can be readily adapted to MCFs with two-dimensional core arrangements. Its implementation will facilitate the development of compact and cost-effective multi-dimensional MDMA systems.

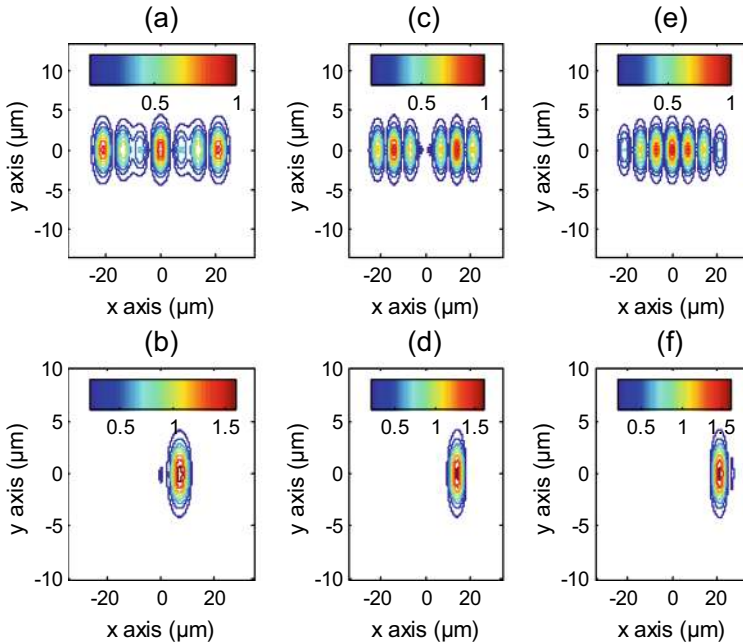


Fig. 6.4 The super-modes of the MCF and the fields at the output of the DEMUX with $D = 7 \mu\text{m}$, **a** the 5th mode, **b** the field at the output of the DEMUX for the 5th mode, **c** the 6th mode, **d** the field at the output of the DEMUX for the 6th mode, **e** the 7th mode, **f** the field at the output of the DEMUX for the 7th mode. This figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1]

6.2 OAM States Generation/Detection Based on the Multimode Interference Effect in a Ring Core Fiber

Part of the section reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18].

In this section, we introduce a proposal to generate and detect orbital angular momentum (OAM) states by using the multimode interference (MMI) effect within a segment of ring core fiber. We present a comprehensive theory that explains the MMI process occurring within the ring core fibers. This MMI process has the capability to transform a single input image into multiple and evenly spaced replicated output images. Following phase adjustments made by a fixed phase shifter array, these output images have the potential to excite OAM states either within the ring core fiber itself or in free space. Consequently, this approach offers a novel method for realizing an OAM multiplexer/de-multiplexer (MUX/DEMUX) through the combination of a ring core fiber segment and a fixed phase shifter array.

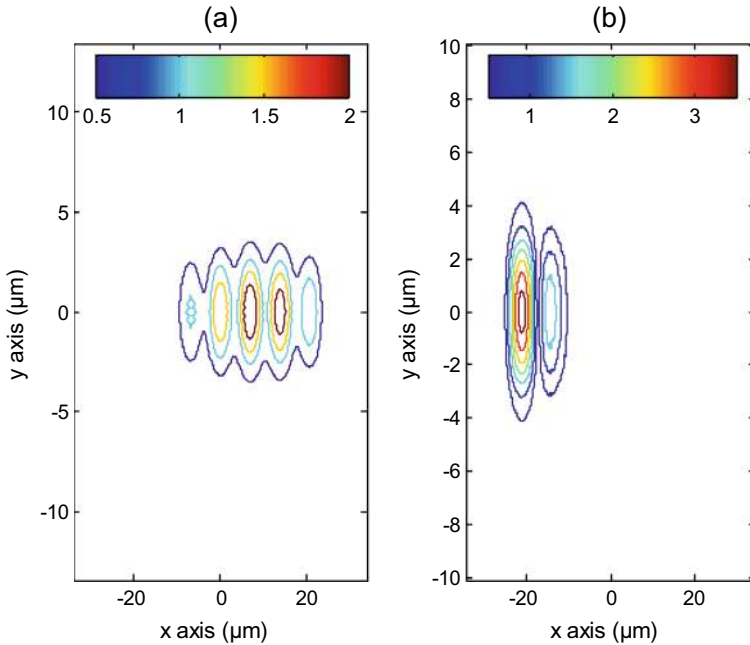


Fig. 6.5 The combination of super-modes of the MCF and the fields at the output of the DEMUX with $D = 7 \mu\text{m}$, **a** the combination of the 1st and the 2nd mode, **b** the field at the output of the DEMUX. This figure is reprinted/adapted with permission from [IEEE Photon. Technol. Lett. 25(13), 1214–1217 (2013)] © IEEE [1]

6.2.1 Introduction

Optical Orbital Angular Momentum (OAM) multiplexing has recently gained significant attention as a potential candidate for spatial division multiplexing (SDM) in optical communication systems, as evidenced by various studies [19–29]. This technique can be applied in both free space [19–25] and ring core fibers [26–29].

The main challenge in implementing this technique lies in the generation and detection of OAM states, whether in ring core fibers or free space. In free space, methods such as spiral phase plates (SPP) [19, 21], dove prisms [22], or ring resonators [23] can be used to generate OAM states. Within ring core fibers, OAM state generation can be achieved through effective index matching and coherent combining [27], transfer of spin angular momentum (SAM) to OAM [28] or by combining a circular array of coherent sources [29]. The latter approach offers scalability for generating numerous OAM modes and integrates seamlessly with multi-core fibers and planar lightwave circuits (PLCs) [24, 25]. This technology can also be adapted to generate OAM states in free space [24, 25]. The principle behind this method involves arranging coherent sources evenly on a circle, with a phase difference of $2\pi l/N$ between adjacent sources. Here, l represents the OAM charge number,

and N denotes the total number of coherent sources in the circular array. Consequently, the phase distribution across the array follows $2\pi nl/N$, where n refers to the n th waveguide in the array. This phase distribution aligns with the spatial sampling of the azimuth phase of the OAM state l . When these coherent sources emit light into a segment of ring core fiber [29] or free space [24, 25], the OAM state l is generated.

While this technique shows promise, it requires pre-coding multiple coherent sources using Discrete Fourier Transform (DFT) [29]. Given that complex DSP algorithms are already implemented in coherent optical communication systems, this requirement increases the demand on digital signal processing (DSP) and makes DSP chip design a critical aspect of system design.

In this section, we introduce a novel method for generating and detecting OAM states. This approach exploits the multimode interference (MMI) effect, which has previously been employed in creating mode multiplexers/de-multiplexers (MUXs/DEMUXs) for planar waveguides [30, 31], rectangular waveguides [32] and coupled waveguide arrays [1]. The MMI effect, well-known in planar waveguides for converting a single input image into multiple phase-shifted duplicates, is explored here in the context of ring core fibers. We establish a comprehensive analytical framework to explain this phenomenon. Within a ring core fiber, the MMI effect similarly transforms an input image into multiple output images with varying phases. By adjusting these phases using a phase shifter array, we can achieve a phase distribution of $2\pi nl/N$, where n represents the n th output port and l is an integer determined by the input port number. Therefore, by combining a ring-shaped MMI coupler with a fixed phase shifter array at the output, we can generate different OAM states by simply injecting the signal at various input ports. Meanwhile, when used in the reverse direction, this device can facilitate OAM state detection. Compared to other OAM MUXs/DEMUXs, our proposed device aligns with the existing optical fiber technology. It offers the versatility to generate multiple OAM states and can be easily scaled up to produce a large number of such states.

6.2.2 MMI Theory in Ring Core Fibers

To begin with, we will derive the MMI theory within a ring core fiber. The mode field inside this fiber can be characterized by the following equation [29]

$$\text{OAM}_{l,p} = R_p(r) \exp(jl\varphi) \exp(j\beta_{l,p}z) \quad (6.14)$$

where p denotes the number of peaks in the radial direction, l represents the OAM charge number in the azimuth domain, and $\beta_{l,p}$ is the propagation constant in the z -direction. It's important to note that although the radial field distribution $R_p(r)$ technically depends on l , if the ring's radius is significantly larger than its width, we can assume that it is independent of l . The ring can be designed to be single-mode in the radial direction, thus making $p = 1$ [29]

$$\text{OAM}_l = R_l(r) \exp(jl\varphi) \exp(j\beta_l z) \quad (6.15)$$

One may reformulate the radial direction distribution as

$$R(r) = R_l(r) \quad (6.16)$$

Substituting Eq. (6.15) into the polar coordinates Helmholtz equation yields:

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial R(r)}{\partial r} \right) - \frac{l^2}{r^2} R(r) + (k^2 - \beta_l^2) R(r) = 0 \quad (6.17)$$

where $k = nk_0$, n is the refractive index of the ring core, and k_0 is the free space wave number. When the ring radius a is relatively large compared to its width, we can approximate that $r \approx a$ within the ring. Hence, Eq. (6.17) can be formulated as

$$R''(r) + \frac{1}{a} R'(r) + \left(k^2 - \beta_l^2 - \frac{l^2}{a^2} \right) R(r) = 0 \quad (6.18)$$

Defining

$$\frac{R''(r) + \frac{1}{a} R'(r)}{R(r)} = k_r^2 \quad (6.19)$$

We can calculate the mode propagation constant as:

$$\begin{aligned} \beta_l &= \sqrt{k^2 - \left(\frac{l^2}{a^2} + k_r^2 \right)} \\ &\approx k \left(1 - \frac{\frac{l^2}{a^2} + k_r^2}{2k^2} \right) \end{aligned} \quad (6.20)$$

The optical field in the ring can be represented by the summation of the OAM modes. Ignoring the common phase term in Eq. (6.20), we may write the optical field as

$$E(r, \varphi, z) = \sum_l R(r) c_l \exp(jl\varphi) \exp\left(-\frac{l^2 z}{2ka^2}\right) \quad (6.21)$$

where c_l denotes the OAM mode coefficient. The MMI effect, which transforms a single input image into N output images, occurs when the length of the ring core fiber satisfies the following condition:

$$L_N = \frac{2ka^2\pi}{N} \quad (6.22)$$

Under this condition, Eq. (6.21) turns to:

$$E(r, \varphi, L_N) = R(r) \sum_l c_l \exp(jl\varphi) \exp\left(-j\frac{l^2\pi}{N}\right) \quad (6.23)$$

Assuming that N uniformly spaced inputs are injected into the ring core fiber at $z = 0$, and the input field distribution in the azimuth domain consists of ideal Dirac functions (we'll consider non-Dirac function cases later), the field at $z = 0$ is given by

$$R(r) \sum_{m=0}^{N-1} a_m \delta\left(\varphi - \frac{2m\pi}{N}\right) \quad (6.24)$$

where a_m stands for the coefficient of each input. By comparing Eqs. (6.21) and (6.24), we can determine the OAM mode coefficient c_l as:

$$c_l = \frac{1}{2\pi} \sum_{m=0}^{N-1} a_m \exp\left(-j\frac{2ml\pi}{N}\right) \quad (6.25)$$

As such, N output images will have their amplitude being the linear superposition of the input coefficients. The relationship between the input and output coefficients is derived as follows.

At $z = L_N$, the optical field should be formulated by substituting Eq. (6.25) into Eq. (6.23). Letting $l = nN + k$, where n is an integer, k an integer between 0 and $N-1$, we may have

$$\begin{aligned} E(r, \varphi, L_N) = & R(r) \frac{1}{2\pi} \sum_n \exp(jnN\varphi) \sum_{k=0}^{N-1} \sum_{m=0}^{N-1} a_m \\ & \exp\left(-j\frac{2mk\pi}{N}\right) \exp\left(-j\frac{k^2\pi}{N}\right) \exp(jk\varphi) \end{aligned} \quad (6.26)$$

Using the formulas below:

$$\begin{aligned} \frac{N}{2\pi} \sum_n \exp(jnN\varphi) &= \sum_{n=0}^{N-1} \delta\left(\varphi - \frac{2\pi n}{N}\right) \quad \varphi \in [0, 2\pi) \\ \sum_{k=0}^{N-1} \exp\left(-j\frac{k^2\pi}{N}\right) &= \sqrt{N} \exp\left(-j\frac{\pi}{4}\right) \end{aligned} \quad (6.27)$$

we can rewrite Eq. (6.26) as

$$E(r, \varphi, L_N) = R(r) \frac{\exp(-j\frac{\pi}{4})}{\sqrt{N}} \sum_{n=0}^{N-1} \left(\sum_{m=0}^{N-1} a_m \exp\left(j\frac{m^2\pi}{N}\right) \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{n^2\pi}{N}\right) \right) \delta\left(\varphi - \frac{2\pi n}{N}\right) \quad (6.28)$$

It is evident from Eq. (6.28) that with N evenly spaced inputs at $z = 0$, there will be N evenly spaced outputs at $z = L_N$. The coefficients of these outputs will be a linear superposition of the input coefficients.

Now, let's consider that the input image with an azimuth distribution denoted by $f(\varphi)$. We may express the input optical field as

$$\begin{aligned} E(r, \varphi, 0) &= R(r) \left(f(\varphi) \otimes \sum_{m=0}^{N-1} a_m \delta\left(\varphi - \frac{2m\pi}{N}\right) \right) \\ &= R(r) \sum_{m=0}^{N-1} a_m f\left(\varphi - \frac{2m\pi}{N}\right) \end{aligned} \quad (6.29)$$

where $\otimes$ denotes convolution. Due to the fact that the MMI process is a linear process, the output can be calculated as

$$\begin{aligned} E(r, \varphi, L_N) &= R(r) \sum_{n=0}^{N-1} b_n f\left(\varphi - \frac{2\pi n}{N}\right) \\ b_n &= \frac{\exp(-j\frac{\pi}{4})}{\sqrt{N}} \sum_{m=0}^{N-1} a_m \exp\left(j\frac{m^2\pi}{N}\right) \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{n^2\pi}{N}\right) \end{aligned} \quad (6.30)$$

Hence, when N input images are equally spaced with a specific azimuth distribution, they will produce N equally spaced output images maintaining the same azimuth distribution. The sole alteration will be in the input and output coefficients, which vary accordingly. The input coefficient a_m ($m = 0, \dots, N-1$) and the output confinement b_n ($n = 0, \dots, N-1$) are grouped in the vector form as $\mathbf{a}$ and $\mathbf{b}$ to facilitate the formulation. These vectors are related through the following equation:

$$\mathbf{b} = \mathbf{T}\mathbf{a} \quad (6.31)$$

The transfer matrix $\mathbf{T}$ can be formulated analytically as

$$\begin{aligned} T_{mn} &= \frac{\exp(-j\frac{\pi}{4})}{\sqrt{N}} \exp\left(j\frac{m^2\pi}{N}\right) \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{n^2\pi}{N}\right) \\ m &= 0, \dots, N-1 \\ n &= 0, \dots, N-1 \end{aligned} \quad (6.32)$$

6.2.3 MMI Effect Based OAM MUX and DEMUX

Based on the analysis presented in the preceding section, if we introduce a phase shifter to the n th output port:

$$\exp\left(-j\frac{n^2\pi}{N}\right) \quad (6.33)$$

Equation (6.31) can be converted to:

$$\begin{aligned} \mathbf{b} &= \mathbf{T}'\mathbf{a} \\ \mathbf{T}' &= \mathbf{D}\mathbf{T} \\ \mathbf{D} &= \begin{pmatrix} \exp\left(-j\frac{0^2\pi}{N}\right) & & & \\ & \ddots & & \\ & & \exp\left(-j\frac{(N-2)^2\pi}{N}\right) & \\ & & & \exp\left(-j\frac{(N-1)^2\pi}{N}\right) \end{pmatrix} \end{aligned} \quad (6.34)$$

$\mathbf{T}'$ will be changed to:

$$\mathbf{T}'_{mn} = \frac{\exp\left(-j\frac{\pi}{4}\right)}{\sqrt{N}} \exp\left(-j\frac{2mn\pi}{N}\right) \exp\left(j\frac{m^2\pi}{N}\right) \quad (6.35)$$

Therefore, in case the m th input port has an input, the phase distribution at the N output ports can be formulated as $-2\pi mn/N$ by ignoring the constant phase term, which can be implemented for OAM state $-m$ generation. Meanwhile, the device can be used to detect OAM state $-m$ at the receiver side by placing the device inversely. Hence, we may realize an OAM MUX/DEMUX with a ring core fiber section whose length is L_N , together with a phase shifter array with the phase shift of $-n^2\pi/N$.

6.2.4 Numerical Simulations

Numerical simulations, utilizing the beam propagation method (BPM), have been conducted to validate the proposed formulas and the OAM MUX/DEMUX concept introduced in previous paragraphs.

A ring core fiber, with parameters similar to those described in reference [29], serves as the MMI device. This ring core fiber features a background index of 1.46. Unlike the 0.12% difference mentioned in [29], the refractive index difference between the core and cladding has been adjusted to 1.2%. This adjustment aims to support a broader range of azimuth modes, facilitating the MMI process (our design accommodates over 45 modes). As is suggested by the theory, an infinite number of

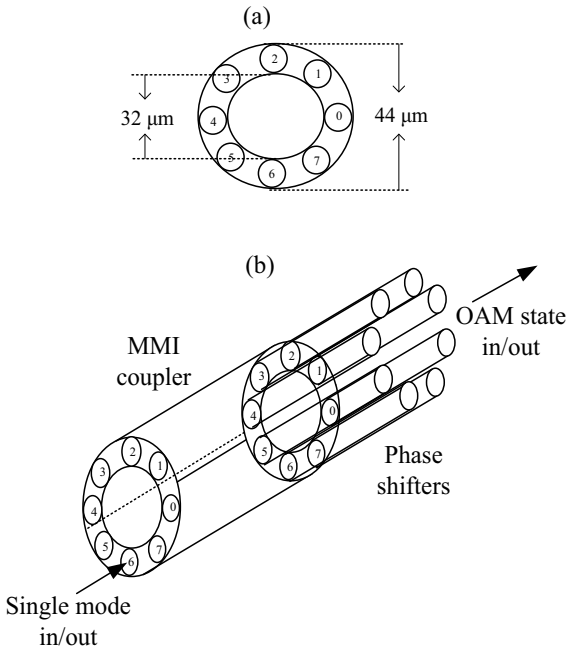
azimuth modes would lead to ideal MMI. Thus, a higher number of azimuth modes enhance performance. It is important to note that our proposed device operates within the weakly guiding regime, making the polarization effect, as is discussed in [33, 34], negligible.

The ring’s dimensions remain unchanged from [29], with an inner radius of 16 μm and an outer radius of 22 μm . Eight ports are evenly distributed at both the input and output. These ports connect to single-mode fibers with a radius of 3 μm , ensuring compatibility between the ring core fiber’s width and the single-mode fibers’ diameters. Operating at a signal wavelength of 1550 nm, the length of the ring core fiber is calculated using Eq. (6.22) and fine-tuned to 1748 μm based on simulation optimizations. It is assumed that the output ports are equipped with phase shifters before the output optical field is used for OAM state generation. Figure 6.6a illustrates the device’s cross-section, while Fig. 6.6b provides a conceptual representation of the device.

Phase shifters capable of imparting a phase shift of $-n2\pi/N$ can be implemented through various methods, including bending array waveguides, waveguides with varying refractive indexes [6], and external phase shifter arrays [16] In our simulations, we have chosen waveguides with distinctive refractive indexes [6] that can deliver the desired phase shift. These waveguides also have a radius of 3 μm .

Initially, let’s demonstrate how a single input image at the MMI device generates N images at the output. Figure 6.7 shows the amplitude and phase distributions at the output of the ring-shaped MMI coupler when input signals are fed into the 0th ($m =$

Fig. 6.6 **a** Cross section of the device with input/output port labeling **b** the conceptual illustration of the device as a mode MUX/ DEMUX. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]



0) and 1st ($m = 1$) input ports. For clarity, the amplitudes in the figures are normalized to a unit peak amplitude (this normalization applies to amplitudes in other figures as well). The figures reveal that a single input image transforms into eight replicas at the device's output. The phase distributions align with the theoretical predictions outlined in Eq. (6.32), disregarding the common phase term. A quantitative assessment of the phases will be provided in the subsequent paragraphs.

Subsequently, the output signals transmit through a phase shifter array consisting of waveguides with varying refractive indexes. Tables 6.1 and 6.2 demonstrate the phases at the center of the eight images, both before and after phase adjustment, when the input signals are fed into the 0th ($m = 0$) and 1st ($m = 1$) input ports, respectively. The tables clearly indicate that, after phase adjustment via the phase shifter array, the phase of each image aligns perfectly with the ideal phase distribution of $-2\pi mn/N$. Consequently, these images can facilitate the generation of OAM state 0 and OAM state -1 , whether within a ring core fiber or in free space.

To illustrate the OAM state generation capability, the output field of the MMI device after phase adjustment by the phase shifter array is fed into a ring core fiber sharing the identical parameters with the one described in [29]. Specifically, the background refractive index is set at 1.46, and the index difference between the

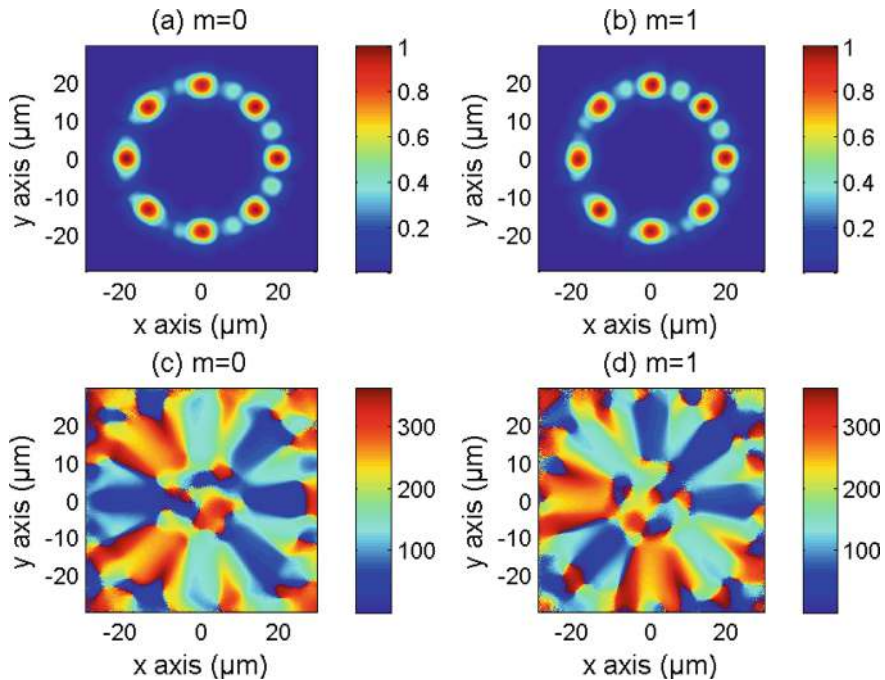


Fig. 6.7 **a, c** The amplitude and phase distribution when the input signal is at the 0th port. **b, d** The amplitude and phase distribution when the input signal is at the 1st port. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

Table. 6.1 Phase distributions of the duplicated images before and after phase adjustment when the input signal is injected into the 0th input port ($m = 0$)

port No	Phase at the output of the MMI coupler	Phase after passing the phase shifters	Ideal phase shift – $2\pi mn/N$
0	0.0	0.0	0
1	22.0	−0.5	0
2	86.8	−3.2	0
3	197.2	−5.3	0
4	−7.7	−7.7	0
5	197.0	−5.5	0
6	86.9	−3.1	0
7	22.4	−0.1	0

This table is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

Table. 6.2 Phase distributions of the duplicated images before and after phase adjustment when the input signal is injected into the 1st input port ($m = 1$)

Port No	Phase at the output of the MMI coupler	Phase after passing the phase shifters	Ideal phase shift – $2\pi mn/N$
0	0.0	0.0	0
1	−21.6	−44.1	−45
2	0.0	−90.0	−90
3	65.8	−136.7	−135
4	174.7	−185.3	−180
5	−29.3	−231.8	−225
6	174.6	−275.4	−270
7	65.9	−316.6	−315

This table is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

core and the cladding is 0.12%. The inner and outer radii are maintained at 16 μm and 22 μm , respectively [29]. As is indicated in [29], such a ring core fiber has the capacity to sustain OAM states ranging from -3 to 3 within the core region, while higher-order OAM modes dissipate into the cladding [29]. Therefore, when optical signals are introduced into the 0th–3rd ($m = 0\text{--}3$) and 5th to 7th ($m = 5\text{--}7$) input ports, the output field of the proposed MUX/DEMUX can trigger OAM states -3 to 3 inside the ring core fiber. However, injecting the signal into the 4th input port of the device leads to the generation of OAM state 4 (combined with OAM state -4), which diminishes after propagation within the ring core fiber.

The OAM states produced within the aforementioned ring core fiber are demonstrated in Figs. 6.8 and 6.9. Corresponding signals are injected into the 0th–3rd ports

($m = 0-3$) of the ring-shaped MMI coupler. Phases are adjusted at the MMI device's output using the phase shifter array, followed by injecting the optical waves into the 8 cm-long ring core fiber. Higher-order modes undergo significant attenuation along the propagation length (8 cm), leaving only OAM states from -3 to 3 persisting within the core. As is evident from Figs. 6.8 and 6.9, OAM states 0 to -3 are generated within the core region while injecting signals into the 0th–3rd input ports ($m = 0$ to 3) of the device, thereby affirming its mode generation proficiency.

To illustrate the OAM state generation process in free space, the output field of the device is projected into the free space and allowed to propagate for 1 cm. Given that the optical field naturally disperses during free-space propagation, the necessary simulation area in both the x and y directions is considerably extensive. Owing to the substantial computational demands, the optical pattern is determined based on Fresnel diffraction theory, rather than BPM. The output amplitude and phase of the generated OAM states in free space are presented in Fig. 6.10. By examining the figures, it becomes evident that OAM state 0 and OAM state -1 are successfully generated when the input signals are introduced into the 0th ($m = 0$) and 1st ($m = 1$)

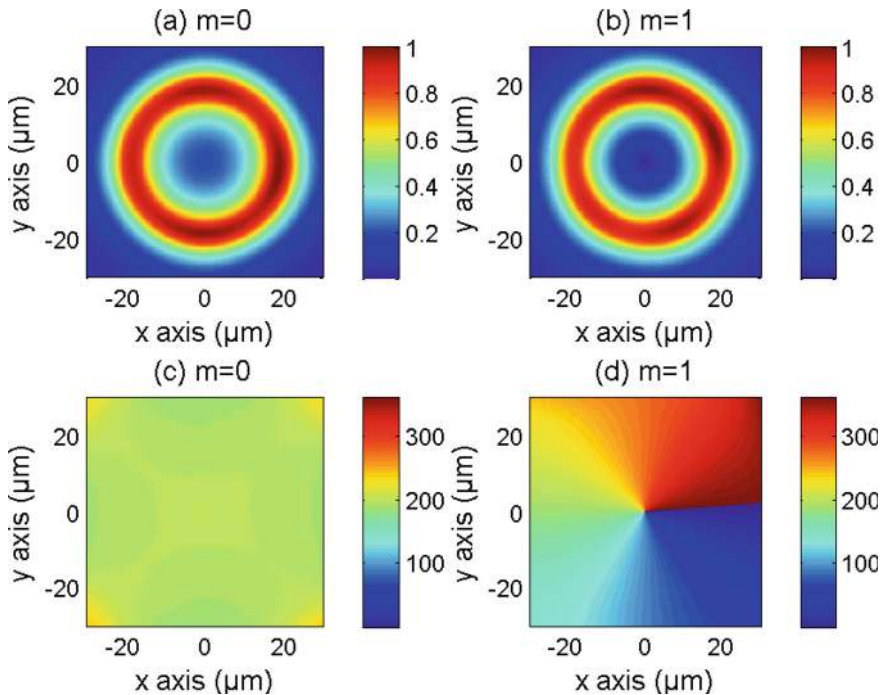


Fig. 6.8 OAM modes generated inside the ring core fiber, **a, c** amplitude and phase of OAM state 0 generated by injecting the signal into the 0th input port of the OAM MUX/DEMUX. **b, d** Amplitude and phase of OAM state -1 generated by injecting the signal into the 1st input port of the OAM MUX/DEMUX. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

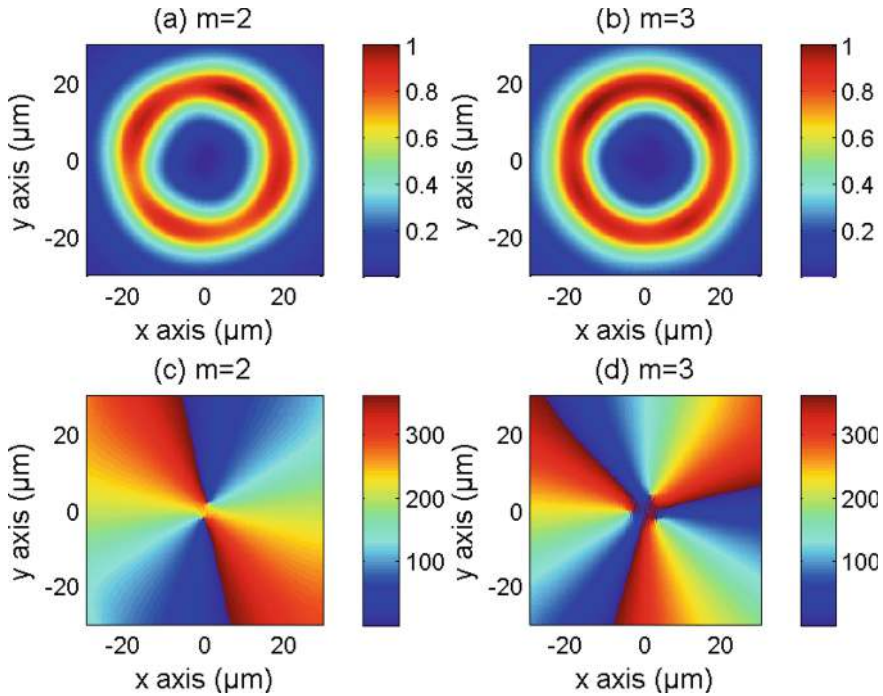


Fig. 6.9 OAM modes generated inside the ring core fiber, **a, c** amplitude and phase of OAM state -2 generated by injecting the signal into the 2nd input port of the OAM MUX/DEMUX. **b, d** amplitude and phase of OAM state -3 generated by injecting the signal into the 3rd input port of the OAM MUX/DEMUX. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

input ports of the MUX/DEMUX, respectively. Consequently, this MUX/DEMUX can be effectively utilized for generating OAM states in free space.

Following the illustration of OAM state generation with our device, we proceed to illustrate OAM state detection. We inversely utilize the device by introducing multiple inputs carrying distinct phases. These input signals represent spatial samples of the OAM states, and the sampling is accomplished through an array of waveguides arranged in a circular pattern, each with a radius of 3 μm . Subsequently, the sampled signals encounter an array of phase shifters, which tune the phase based on Eq. (6.33), before experiencing the MMI process within the ring core fiber.

During the de-multiplexing process, all parameters for the MMI device and the phase shifters keep unaltered. We use OAM states within a ring core fiber characterized by a background index of 1.46 and a 1.2% index difference between the core and cladding in our de-multiplexing simulations. The reason for selecting this particular ring core fiber, instead of the one referenced in [29], is that the latter cannot sustain the OAM state 4(-4) within its structure.

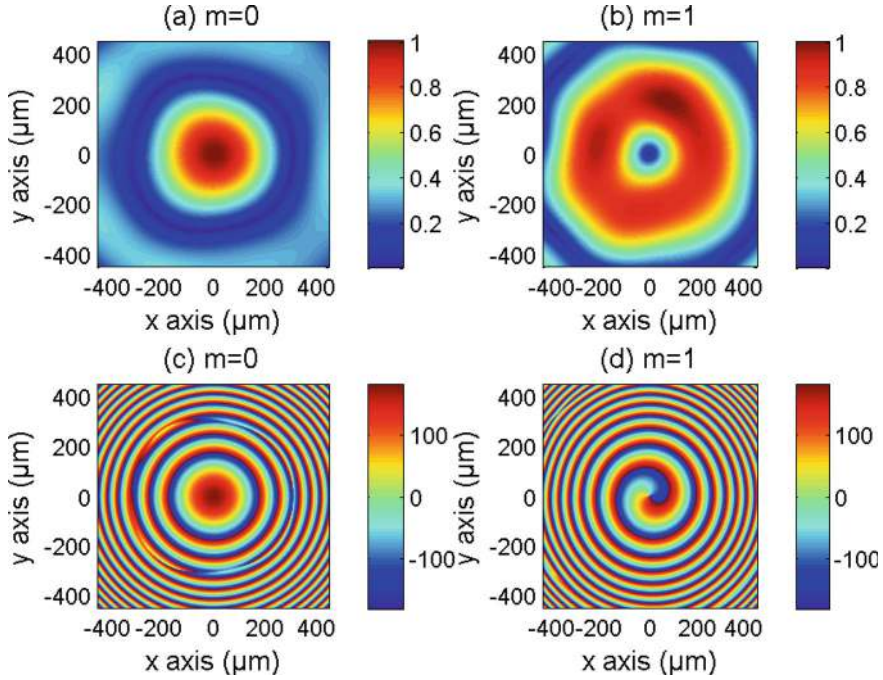


Fig. 6.10 OAM modes generated in free space, **a, c** amplitude and phase of OAM state 0 generated by injecting the signal into the 0th input port of the OAM MUX/DEMUX. **b, d** amplitude and phase of OAM state -1 generated by injecting the signal into the 1st input port of the OAM MUX/DEMUX. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

We spatially sample OAM states ranging from -3 to 3 , as well as OAM state 4 , and inversely inject them into our proposed MUX/DEMUX. The outputs are visualized in Figs. 6.11 and 6.12. Depending on the input OAM state, the output optical field focuses on specific output ports, with the port number corresponding to the input OAM state. Specifically, OAM states $0, -1, -2, -3$, and $4(-4)$ (denoted by $m = 0$ to 4) yield outputs at ports $0, 1, 2, 3$, and 4 respectively, while OAM states $3, 2, 1$ (denoted by $m = 5$ to 7) produce outputs at ports $5, 6$, and 7 .

As is evident from the figures, our proposed device successfully de-multiplexes the OAM states. The mode extinction ratio, defined as the ratio of optical power at the designated output port to the power distributed across other ports, maintains a high value of 29 dB. This signifies that crosstalk from one mode to others remains below -29 dB, indicating robust performance.

At last, a tolerance analysis for fabrication errors was conducted. Since both OAM state generation and detection are facilitated by the same device, we assessed the impact of manufacturing imperfections on OAM state detection as a representative case. In scenarios where the length of the ring core fiber experiences a 2% error, equivalent to approximately $34\text{ }\mu\text{m}$, the mode extinction ratio diminishes to

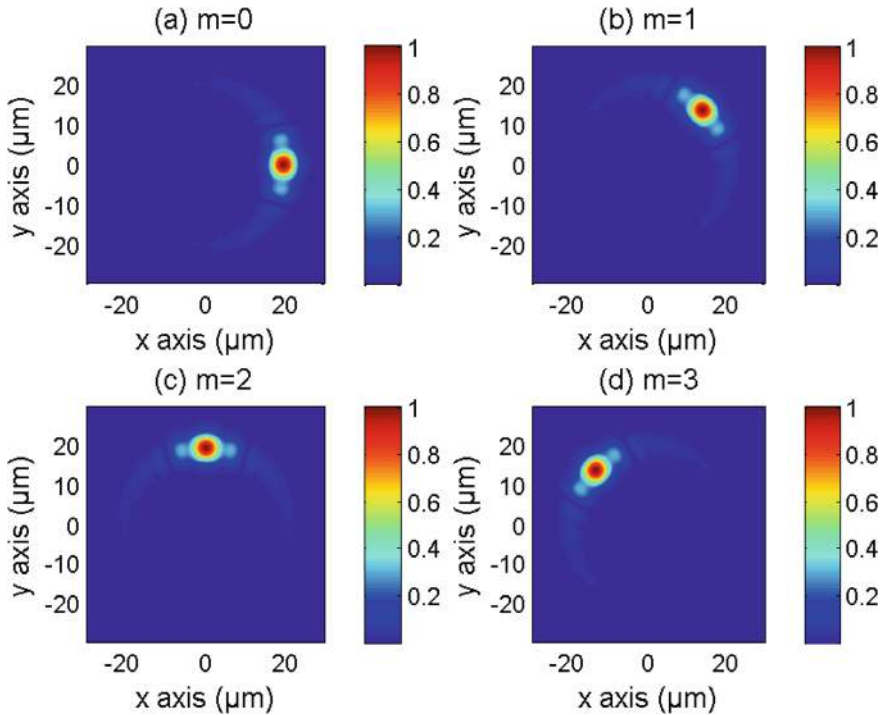


Fig. 6.11 OAM state detection using the device. **a–d** OAM states 0, -1 , -2 , and -3 ($m = 0, 1, 2$, and 3) are spatially sampled and passed onto the device, and figures show the field patterns at the output of the device. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

25 dB. Similarly, when a single phase shifter encounters a phase deviation of 20 degrees, the mode extinction ratio drops to 24 dB. Alterations in the radius of the ring core fiber can be correlated with length variations through Eq. (6.22), indicating that a 1% fluctuation in the average radius of the ring corresponds to a 2% change in length. These simulations reveal that the device exhibits considerable resilience against geometrical deviations. Nevertheless, phase discrepancies can significantly undermine the performance. Hence, it is advisable to incorporate an adjustable phase shifter array at the device's output.

The device operational wavelength range can be approximately obtained. Based on Eq. (6.22), deviations in wavelength can be proportionally linked to deviations in length. Consequently, with a $\pm 2\%$ fluctuation in signal wavelength (± 30 nm) centered around 1550 nm, the DEMUX keeps a mode extinction ratio of 25 dB.

Last but not the least, the influence of random variations in the ring's radius was explored. We assume random variations with the standard deviations of $0.5 \mu\text{m}$ in both the inner and outer ring radii, while maintaining constant average radii. Such deviation is considerable from a fiber technology perspective. Even under such

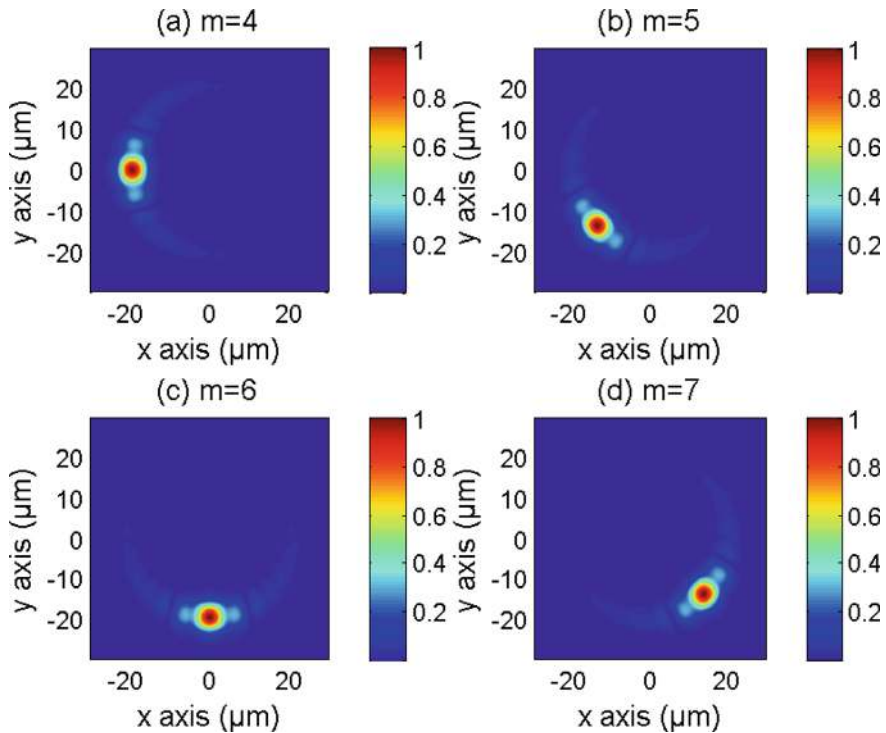


Fig. 6.12 OAM state detection using the device. **a–d** OAM states 4, 3, 2, and 1 ($m = 4, 5, 6$, and 7) are spatially sampled and passed onto the device, and the figures show the field patterns at the output of the device. This figure is reprinted/adapted with permission from [Opt. Express 23, 10,247–10,258 (2015)] © The Optical Society of America [18]

extreme conditions, the mode extinction ratio persists at 26 dB. This underscores the structure's robustness against random geometrical fluctuations.

6.2.5 Summary

In conclusion, we have introduced an innovative OAM MUX/DEMUX exploiting the MMI effect within a ring core fiber. A comprehensive theoretical framework has been established for the MMI process in ring core fibers. Based on this theory, it has been found that the OAM MUX/DEMUX can be achieved through a segment of ring core fiber together with an array of fixed phase shifters. Detailed numerical simulations have been presented to validate the proposed theory and the device's capability. This novel MUX/DEMUX aligns with prevailing fiber technology and is applicable for both OAM state generation/detection within ring core fibers and in free space.

6.3 A Tunable Arbitrary Unitary Transformer Based on Multiple Sections of Multicore Fibers with Phase Control

Part of the section reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35].

The previous section discussed fixed mode multiplexers/demultiplexers. In this section, we introduce an innovative adjustable unitary transformer capable of executing any discrete unitary conversions, including the multimode fiber mode transformation. This unitary transformer consists of several segments made up of multicore fibers featuring tightly aligned, coupled cores. To manipulate the phases of the waves within the cores, phase shifters are positioned both preceding and following these segments. We present a straightforward algorithm to determine the optimal phase configuration for these shifters, thereby achieving the intended unitary transformations. Given that our proposed device is fiber-based, it is especially applicable for mode division multiplexing systems. Based on this design, we have demonstrated a tunable mode MUX/DEMUX tailored for a three-mode fiber.

6.3.1 Introduction

Mode division multiplexing (MDM) is a prominent option to enhance the transmission capacity of future optical communication systems [36, 37]. Within this framework, mode multiplexers/demultiplexers (MUX/DEMUXs) [6, 38–44] and mode switches constitute crucial network components, drawing significant attention from the industry. The implementation of mode MUX/DEMUX can be achieved through various methods, including the effective index matching technique [6, 38–40] and the interferometric approach [41–43].

Seamless mode conversion and switching without any loss can be considered as optical unitary transforms. This principle has already been employed in the realization of mode MUX/DEMUXs [44]. Thanks to the inherent advantages of the unitary transformation-based methodology [44], the resultant mode MUX/DEMUXs exhibit notably lower mode-dependent loss (MDL) and insertion loss (IL) when compared to those by alternative methods. Beyond their applications in MDM systems, unitary transforms have found extensive applications in other areas, such as quantum information processing [45, 46]. Consequently, a tunable and versatile unitary transformer, capable of facilitating mode conversion and switching across various multimode waveguides while enabling arbitrary photon interference, holds particular appeal for the research community.

The implementation of a tunable arbitrary unitary transformer poses significant challenges. A commonly adopted technique is the multi-plane method, which involves the manipulation of spatial fields through Fourier transforms facilitated by lenses, and the alteration of phases between lenses via holograms. In an effort to

simplify the system and minimize insertion loss, G. Labroille et al. [46] introduced a reflection and feedback approach, utilizing the multi-plane concept for mode MUX-DEMUX realization. This groundbreaking work has laid foundation for developing an arbitrary optical unitary transformer. However, it remains a tailored and static solution, lacking ability to adapt to the varying waveguide configurations or to switch the waveguide modes. While the spatial light modulator (SLM) [47] offers phase tuning flexibility, it is not economically feasible for widespread communication system deployment. Seeking a cost-effective tuning solution, Tang [48] presented a tunable arbitrary unitary transformer exploiting the multi-plane concept in a multi-mode interference coupler. Despite these advancements, there is still a demand for fiber-based tunable arbitrary optical unitary transformers. This is because free-space optics-based [44, 47] or waveguide-based devices [48] induce significant insertion loss when integrated into fiber-related applications. Furthermore, the phase tuning algorithm merits exploration, as the randomized optimization techniques employed in [47, 48] may result in time-consuming system implementations. An alternative approach involves the use of cascaded tunable couplers and phase shifters [49–51], which provides enhanced tunability. Nevertheless, this method requires a considerable number of tunable couplers and may lead to increased loss as the number of cascading stages rises.

In this study, we introduce a novel fiber-based approach for realizing an adjustable arbitrary unitary transformer. We utilize multiple segments of multi-core fibers (MCFs) in conjunction with phase controllers positioned before and after them. The cores of the MCFs are tightly aligned, enabling core-to-core coupling. Consequently, the input signal undergoes transformation within the MCF. By incorporating phase controllers, we can mimic the multi-plane concept. We propose a straightforward yet effective method for adjusting the phase shifters. Numerical simulations validate our proposed concept, and we have designed and simulated a tunable mode MUX/DEMUX tailored for a multimode fiber based on this approach.

6.3.2 Theory

6.3.2.1 Transformation for Each Section of the MCFs

There exist two distinct varieties of MCFs. The first variety features cores that are mutually isolated, whereas the second variety involves coupling cores. Although most MCF transmission system applications prefer the former, our present study chooses for the latter.

Without compromising generality, let us presume that the cores are arranged in a circular formation [52]. During propagation, the amplitudes of waves within these cores follow the coupled equation:

$$\frac{da_1}{dz} = i\beta a_1 + iCa_N + iCa_2$$

$$\begin{aligned}
\frac{da_i}{dz} &= i\beta a_i + iCa_{i-1} + iCa_{i+1} \\
\frac{da_N}{dz} &= i\beta a_1 + iCa_{N-1} + iCa_1
\end{aligned} \tag{6.36}$$

where a_i denotes the amplitude of wave in the i th core, N represents the total count of cores and C signifies the coupling coefficient between adjacent cores. By reformulating Eq. (6.36) into a matrix representation, one arrives at:

$$\begin{aligned}
\frac{d\mathbf{a}}{dz} &= i\beta\mathbf{I}\mathbf{a} + i\mathbf{M}\mathbf{a} \\
\mathbf{M} &= \begin{pmatrix} 0 & C & 0 & \dots & 0 & C \\ C & 0 & C & 0 & \dots & 0 \\ 0 & C & 0 & C & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \ddots & \ddots & \ddots & C \\ C & 0 & \dots & 0 & C & 0 \end{pmatrix}
\end{aligned} \tag{6.37}$$

Matrix $\mathbf{M}$ is a circulant matrix, which is able to be decomposed as [46]

$$\begin{aligned}
\mathbf{M} &= \mathbf{F} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_N \end{pmatrix} \mathbf{F}^H \\
F_{mn} &= \frac{1}{\sqrt{N}} \exp\left(-j\frac{2\pi}{N}(m-1)(n-1)\right) \\
\lambda_k &= 2C \cos\left(\frac{2\pi(k-1)}{N}\right)
\end{aligned} \tag{6.38}$$

where $\mathbf{F}$ is the Fourier transform matrix. Therefore, one may obtain the solution to Eq. (6.37) as [52]:

$$\begin{aligned}
\mathbf{a}(z) &= e^{j\beta z} \mathbf{F} \mathbf{D} \mathbf{F}^H \mathbf{a}(0) \\
\mathbf{D}(z) &= \begin{pmatrix} \exp(\lambda_1 z) & & & \\ & \exp(\lambda_2 z) & & \\ & & \ddots & \\ & & & \exp(\lambda_N z) \end{pmatrix}
\end{aligned} \tag{6.39}$$

6.3.2.2 Arbitrary Unitary Transformer

Assuming K sections of multi-core fibers are connected in series, with the length of the k th segment denoted as L_k . To achieve arbitrary unitary transformations, we emulate the multi-plane approach and introduce phase controllers prior to each MCF segment. The phase controller array preceding the k th segment contributes as a diagonal matrix, denoted as $\theta^{(k)}$. Consequently, the input and output vectors are related by the following equation:

$$\mathbf{a}_{out} = e^{j\beta \sum_k L_k} \prod_k (\mathbf{F}\mathbf{D}(L_k)\mathbf{F}^H \theta^{(k)}) \mathbf{a}_{in}$$

$$\theta^{(k)} = \begin{pmatrix} \exp(j\theta_1^{(k)}) & & & \\ & \exp(j\theta_2^{(k)}) & & \\ & & \ddots & \\ & & & \exp(j\theta_N^{(k)}) \end{pmatrix} \quad (6.40)$$

When implementing our structure for mode MUX/DEMUX applications, Eq. (6.40) should incorporate a transfer matrix that correlates the amplitudes of modes in the multimode fiber with those of the waves in the MCF cores.

From Eq. (6.40), it becomes evident that the transfer matrix of our proposed structure is similar to the transfer matrix employed in the multi-plane method, which contains a phase shifter matrix between two Fourier transform matrices. To enhance control flexibility, an additional phase shifter array can be installed after the final MCF segment. The comprehensive mathematical derivation presented in 6.5.5 confirms that our proposed structure can indeed realize an arbitrary unitary transformer. Although the derivation in Sect. 6.5.5 suggests a substantial number of MCF couplers and phase shifters are needed, numerical testing verifies that our structure is capable of achieving arbitrary unitary transforms if the total number of phase controllers in summation with the constraints on matrix elements exceeds the number of unknowns in the unitary matrix. When implementing our structure for mode MUX/DEMUX applications, Eq. (6.40) should contain a transfer matrix which correlates the amplitudes of modes in the multimode fiber with those of the waves in the MCF cores.

It is important to note that the use of a circular array is not mandatory. Alternative core array alignments, such as linearly aligned cores, can also facilitate the realization of arbitrary unitary transformations with the assistance of phase shifters.

6.3.2.3 Simple Algorithm to Realize the Optimal Phase Setup

Finding the optimal control parameters for phase shifters can be challenging due to the large number of variables requiring optimization. The optimization problem is neither

linear nor quadratic, making efficient optimization methods like linear programming and quadratic programming inapplicable. Random optimization techniques, such as the simulated annealing algorithm [47, 48] and the genetic algorithms, can be employed to determine the optimal phase configuration. However, their significant computational complexity makes it impractical for phase tuning in real time systems. Gradient search is another option, but it can also be computationally demanding when dealing with a large number of variables.

Hence, a simplified approach for the phase optimization is proposed by sequentially adjusting the phase shifters. We define the cost function as:

$$E = \sum_{m,n} |T_{mn} - O_{mn}|^2 \quad (6.41)$$

where T_{mn} stands for the transfer matrix element, and O_{mn} stands for objective unitary matrix (**O**) element. In each step, we optimize one phase shifter while keeping others altered so that the cost function E_k is minimized. It is straight forward to prove the convergence of E_k . Since

$$0 < \dots E_{k+1} \leq E_k \leq E_{k-1} \dots \quad (6.42)$$

E_k is a descending sequence with a lower bound, which will converge to the limit. We can use this limit as the optimal configuration for the phase shifters. To prevent getting stuck in local optima, the order of phase shifter optimization is randomized during each round. For instance, one might optimize phase shifter 2, then phase shifter 1, followed by phase shifter 3, and so on, during the first optimization round, while adopting a different optimization order in the subsequent round. It is important to note that the proposed method's performance might not match that of global optimization techniques like simulated annealing. However, it offers a balance between efficiency and performance.

When developing devices such as power splitters, mode MUX/DEMUXs, and mode switches based on the arbitrary unitary transformer, there are no specific needs regarding the phases. As such, the cost function can be reformulated as:

$$E = \sum_{m,n} ||T_{mn}| - |O_{mn}||^2 \quad (6.43)$$

6.3.3 Results and Discussions

Numerical simulations have been conducted to validate the proposed device and optimization algorithm. Initially, we assessed the efficacy of our approach in achieving arbitrary unitary transformations. For this, we employed six segments of four-core MCFs, with five sets of phase shifters interposed between each pair of segments.

Table 6.3 The target matrix: the 4×4 DFT matrix

0.5	0.5	0.5	0.5
0.5	$-0.5i$	-0.5	$0.5i$
0.5	-0.5	0.5	-0.5
0.5	$0.5i$	-0.5	$0.5i$

This table is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

Additionally, the first and last MCF segments were linked via two sets of phase shifters, i.e., seven sets of phase shifters were used in total. The MCFs were presumed to possess a coupling strength and length product (CL) of 0.35π (though this value may vary as long as all elements within the transfer matrix are maintained as non-zero). Sequential tuning of the phase shifters was performed to determine the optimal configuration.

Initially, we set the target matrix as the well-known DFT matrix, with its components outlined in Table 6.3. By adjusting the phase shifters in our proposed device with a resolution of 1 degree, we were able to achieve the transfer matrix exhibited in Table 6.4. A comparison between the two tables reveals a remarkable similarity between the matrix elements. The mean square error (MSE) between them is 1.73×10^{-5} .

Furthermore, to evaluate the device's tolerance, we selected a completely random 4×4 matrix, presented in Table 6.5. Once again, by fine-tuning the phase controllers, we could obtain a transfer matrix that closely resembles the target matrix (displayed in Table 6.6). The MSE between these two matrices is 1.85×10^{-4} .

Table 6.4 The realized transfer matrix with Table 6.3 as the target

$0.4948 - 0.0035i$	$0.5011 + 0.0000i$	$0.5010 + 0.0007i$	$0.5031 - 0.0006i$
$0.5006 + 0.0055i$	$0.0028 - 0.5057i$	$-0.4956 - 0.0041i$	$-0.0010 + 0.4981i$
$0.5027 + 0.0004i$	$-0.4964 + 0.0037i$	$0.5015 - 0.0057i$	$-0.4994 - 0.0007i$
$0.5019 + 0.0013i$	$0.0047 + 0.4968i$	$-0.5019 - 0.0008i$	$0.0009 - 0.4995i$

This table is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

Table 6.5 The target matrix: a 4×4 random matrix

$-0.3146 + 0.4636i$	$-0.0736 + 0.4996i$	$0.0432 + 0.4144i$	$-0.2334 + 0.4505i$
$0.2468 - 0.3336i$	$0.4868 - 0.0783i$	$-0.5222 + 0.1347i$	$-0.2915 + 0.4570i$
$-0.5596 + 0.2379i$	$0.3275 + 0.0331i$	$-0.4397 - 0.4080i$	$0.4026 + 0.0073i$
$-0.1821 - 0.3332i$	$0.4617 + 0.4247i$	$0.0981 + 0.4075i$	$0.0630 - 0.5316i$

This table is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

Table 6.6 The realized transfer matrix with Table 6.5 as the target

$-0.3138 + 0.4434i$	$-0.0764 + 0.5194i$	$0.0478 + 0.4191i$	$-0.2305 + 0.4452i$
$0.2439-0.3360i$	$0.4881-0.0755i$	$-0.5218 + 0.1353i$	$-0.2928 + 0.4554i$
$-0.5775 + 0.2387i$	$0.3318 + 0.0307i$	$-0.4359-0.4036i$	$0.3814 + 0.0005i$
$-0.1762-0.3331i$	$0.4548 + 0.4031i$	$0.0915 + 0.4125i$	$0.0555-0.5541i$

This table is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

To assess the proficiency of the arbitrary unitary transformer with randomly chosen target matrices, we computed the MSE between the unitary matrix achieved by the transformer and a randomly selected unitary target matrix. The cumulative distribution function (CDF) of MSE, based on 1000 random realizations, is plotted in Fig. 6.13. As is evident from the figure, the CDF attains 100% when the MSE reaches 3×10^{-3} . These statistical findings demonstrate that the proposed design can accomplish arbitrary unitary transformations with remarkable precision.

Furthermore, we applied our concept to realize a tunable MUX/DEMUX based on a cascaded MCF structure. This MUX/DEMUX was customized for a three-mode fiber. The input modes for this multimode fiber were transformed into MCF modes via a photonic lantern [53, 54]. The schematic for this converter is exhibited in Fig. 6.13,

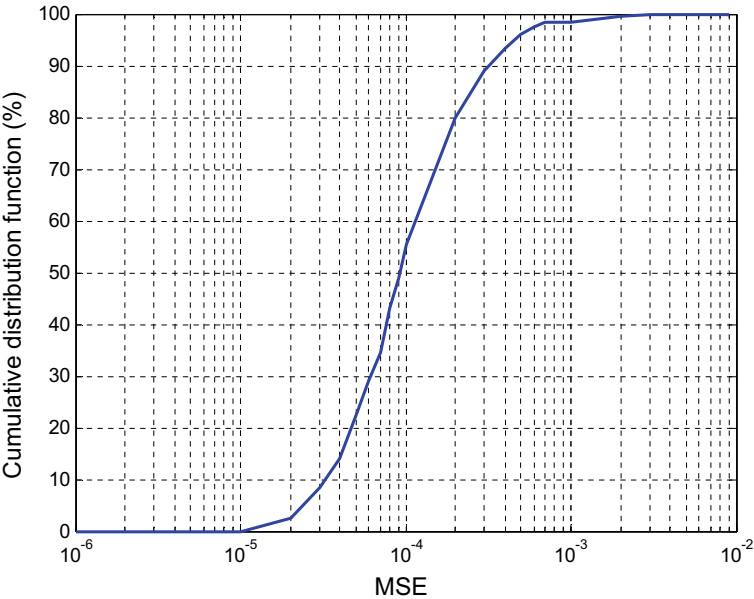


Fig. 6.13 Cumulative distribution function (CDF) of MSE with respect to 1000 random realizations. This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

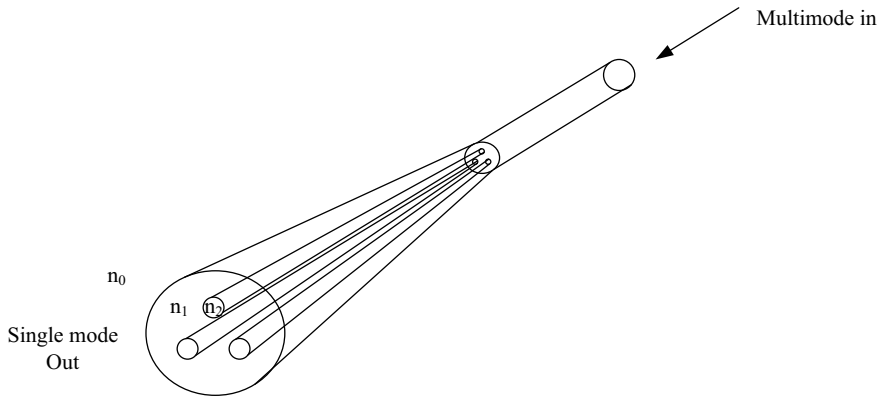


Fig. 6.14 Schematic of the photonic lantern for multimode to single mode conversion. This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

featuring a tapered multi-core fiber structure [54]. The multimode fiber possesses the core index of n_0 and the cladding index of n_1 . Meanwhile, the photonic lantern possesses the outer cladding index of n_0 , the inner cladding index of n_1 and the core index of n_2 , with the values $n_0 = 1.435$, $n_1 = 1.44$ and $n_2 = 1.445$ respectively [53]. Operating at a signal wavelength of 1550 nm, the photonic lantern has a taper length of 20 mm, and its inner cladding radius gradually expands from 6 to 35 μm . The multimode fiber's core radius is 6 μm , while the single-mode cores at the lantern's output have core radii of 4 μm , carefully designed to ensure both single-mode and triple-mode (with two degenerate modes) operations. The core pitch at the lantern's output is set at 41.6 μm , minimizing coupling between cores to a negligible level. The schematic of the photonic lantern is presented in Fig. 6.14. Importantly, the proposed geometric parameters are feasible for implementation in practical fiber/photonic lantern fabrications [53, 55].

Using the photonic lantern technology, it is feasible to transform the modes propagating in a multimode fiber into the modes of a MCF, as is illustrated in Fig. 6.15. By analyzing the results from the BPM simulations depicted in Fig. 6.15, one can compute the value of matrix B that correlates the multimode fiber mode amplitudes to the MCF core amplitudes. The remaining section of the MUX/DEMUX device, implemented via cascading MCFs equipped with phase shifters, aims to obtain a transfer matrix G . The goal is to ensure that the product GB results in a matrix where each column/row contains only one non-zero element.

$$\mathbf{G} = \prod_k \mathbf{K} \theta^{(k)} \quad (6.44)$$

where $\mathbf{K}$ represents the coupler transfer matrix. Through this process, each multimode fiber mode is mapped onto a specific core mode of the MCF.

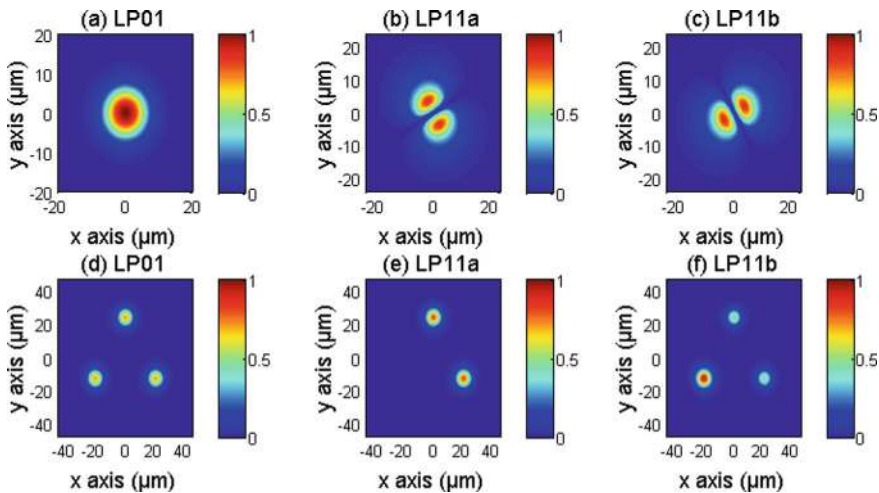


Fig. 6.15 The converted pattern from the modes in the multimode. **a** LP01, **b** LP11a and **c** LP11b to the modes in the multi-core single mode fibers (**d–f**). This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

The MCF-based couplers employed in the tunable MUX, as is depicted in Fig. 6.16, feature cores with an index of 1.445 and a cladding index of 1.44. These couplers have a core radius of 4 μm . At both ends of the section, two tapers are positioned to gradually adjust the core pitch from 20.8 to 41.6 μm , each has a length of 20 mm. The central portion of the MCF-based coupler comprises a strong coupling region, where the core pitch remains constant at 20.8 μm across a length of 17 mm. Depending on the input received from each MCF port, the coupler generates a corresponding output pattern, as is illustrated in Fig. 6.17. Furthermore, we computed the transfer matrix K of the coupler using the data in Fig. 6.17.

The tunable mode MUX/DEMUX, as is illustrated in Fig. 6.18, comprises a photonic lantern, four MCF-based couplers, and four sets of phase shifters, with each set containing three phase shifters. Notably, the relative phase of input/output modes does not affect the unitary transform in this application; thus, phase shifters were omitted after the output MCF couplers. This design choice is justified by the fact that a 3×3 unitary transform involves 13 unknowns and 6 matrix element constraints, requiring a minimum of 7 phase controllers. Since only relative phase is relevant, each phase array effectively has two controllers. Consequently, the use of 4 phase shifter arrays, equivalent to 8 effective controllers, and 5 MCFs is the most compact configuration.

In the simulation process, the photonic lantern and MCF-based couplers were modeled using the BPM, incorporating the structural parameters outlined earlier. Phase shifters were simulated by adjusting the index within the single-mode waveguides used for connection. The overall transfer matrix, linking the multimode fiber mode amplitudes to those of the MCF cores, follows the following equation:

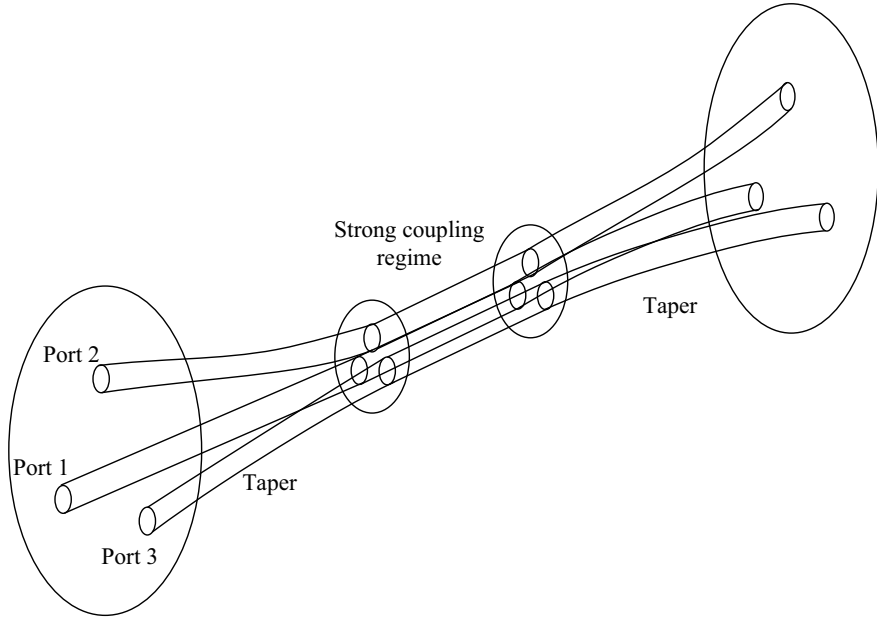


Fig. 6.16 Schematic of the multi-core fiber based three-port coupler. This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

$$\mathbf{T} = \left(\prod_{k=1}^4 \mathbf{K} \theta^{(k)} \right) \mathbf{B} \quad (6.45)$$

Prior to the comprehensive simulation, we determined the transfer matrices for both the photonic lantern $\mathbf{B}$ and the MCF-based coupler $\mathbf{K}$ from numerical data. We then optimized the phase shifter values using a sequential optimization algorithm, aiming for specific target matrices such as $\text{abs}(\mathbf{T}) = [1 \ 0 \ 0; 0 \ 1 \ 0; 0 \ 0 \ 1]$ for mode MUX/DEMUX operations, $\text{abs}(\mathbf{T}) = [0 \ 0 \ 1; 0 \ 1 \ 0; 1 \ 0 \ 0]$ to switch the outputs for mode 1 and mode 3, or other desired configurations. Since output phase doesn't affect performance, we focused on the absolute value of the transfer matrix $\mathbf{T}$ in our optimization.

Using the optimized phases in the final BPM simulations, we achieved perfect de-multiplexing of LP01, LP11a, and LP11b modes to cores 1, 2, and 3 of the MCF, respectively, as is shown in Fig. 6.19. Ignoring the splicing loss and the phase shifter attenuation, the device performance was impressive: over 35 dB mode extinction ratio, less than 0.3 dB insertion loss, and under 0.1 dB MDL. Afterwards, the target matrix $\mathbf{O}$ was adjusted as $[0 \ 0 \ 1; 0 \ 1 \ 0; 1 \ 0 \ 0]$, we successfully switched the output cores for modes LP01 and LP11b as shown in Fig. 6.20, with phase shifter values changed to adapt to the target change. The performance maintains impressive: over 35 dB mode extinction ratio, less than 0.3 dB insertion loss, and less than 0.1 dB MDL.

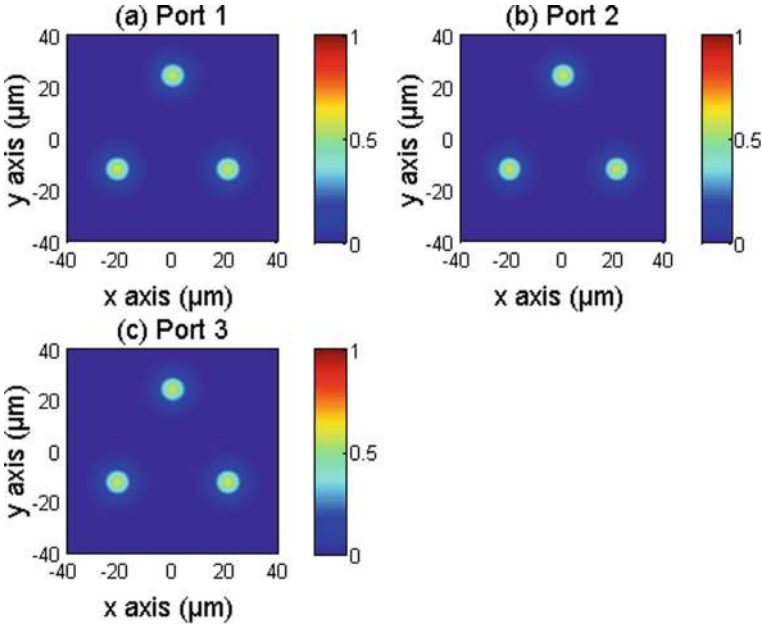


Fig. 6.17 The output pattern of the MCF based coupler with the input from **a** port 1, **b** port 2 and **c** port 3. This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

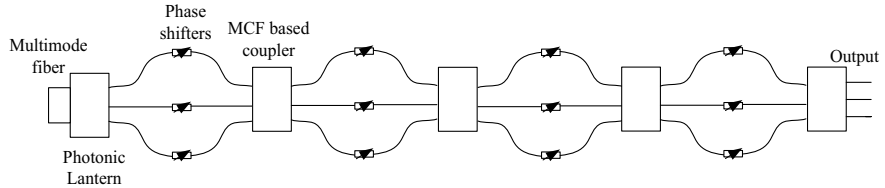


Fig. 6.18 The overall structure of the tunable mode MUX/DEMUX. This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

Further analysis indicates that output core switching for any two modes is achievable without compromising performance, through precise phase shifter adjustments.

The operational bandwidth of the proposed device has been thoroughly examined. The study reveals that, as the central signal wavelength fluctuates by ± 15 nm, ± 10 nm, and ± 5 nm, the insertion loss and MDL remain relatively unaffected. However, there are notable changes in the worst-case mode extinction ratio, shifting to 17 dB, 20 dB, and 26 dB, respectively. From a system implementation viewpoint, the mode extinction ratio stays within acceptable limits throughout the 30 nm-wide C band.

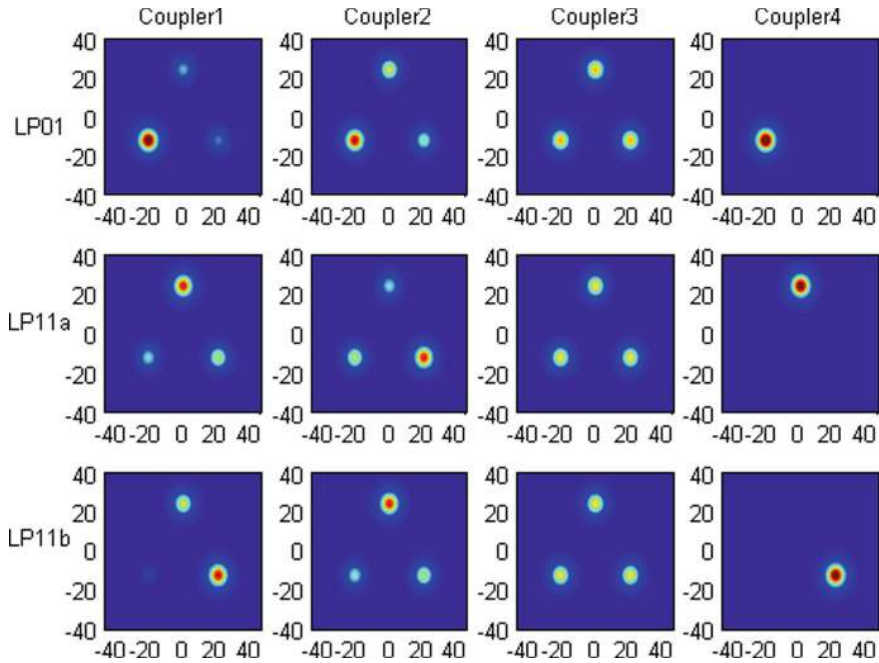


Fig. 6.19 The output of each MCF based coupler with the target transfer matrix $\mathbf{O} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, the x and y axis units are μm . This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

Additionally, a comprehensive robustness analysis has been conducted on the proposed device, taking into account various parameters. When the lengths of the tapers and the coupling conditions for the photonic lantern and MCF-based couplers vary by ± 2 mm, no noticeable decline in mode MUX/DEMUX performance is detected. This resilience can be attributed to the phase shifters' capability to self-adjust, accommodating such parameter deviations.

The study also takes into account the unbalanced coupling arising from variations in core pitch and core diameter, as is documented in references [56, 57]. Despite fluctuations of ± 1 μm in the center positions and ± 0.1 μm in diameters of the MCF cores, no deterioration in performance is observed. Additionally, we substituted the MCF coupler proposed here with a practical 3×3 fiber coupler, utilizing the experimentally measured transfer matrix from [58]. This realistic coupler exhibits loss and an uneven distribution of power among its three output ports. Even with this coupler, the mode extinction ratio surpasses 29 dB, the insertion loss remains under 1.6 dB, and the MDL is less than 0.4 dB.

Furthermore, we examined the device's polarization dependence. Assuming a 0.1 dB PDL and a 5-degree polarization-dependent phase error in our simulations, the mode extinction ratio stays above 25 dB, which is still relatively high and suitable for practical applications.

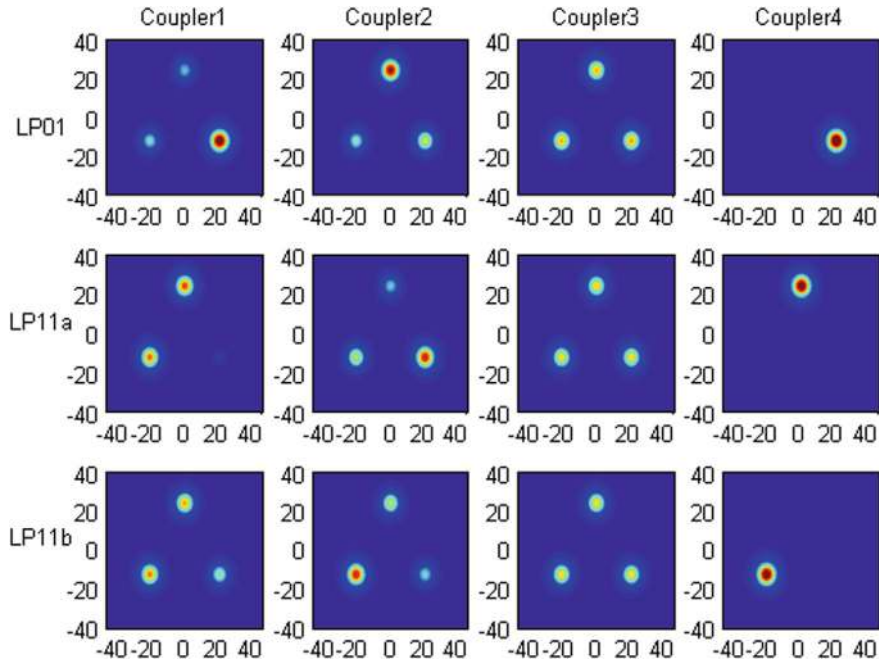


Fig. 6.20 The output of each MCF based coupler with the target transfer matrix $\mathbf{O} = \begin{bmatrix} 0 & 0 & 1; & 0 & 1 & 0; \\ 1 & 0 & 0; \end{bmatrix}$, the x and y axis units are μm . This figure is reprinted/adapted with permission from [Opt. Express 26, 3020–3036 (2018)] © The Optical Society of America [35]

In our final analysis, we incorporated the core-dependent insertion loss resulting from the photonic lantern, phase shifters, and the splicing process, factors that can affect the unitary nature of the transfer matrix. As reported in [59], the core-dependent loss is less than 0.15 dB, while the overall insertion loss, including the core-dependent component, is less than 0.25 dB (with a maximum of 0.4 dB). Adopting these values, we assumed a 0.5 dB core-dependent insertion loss for each stage, accounting for contributions from the MCF coupler, phase shifters, and splicing process. Our findings reveal that the worst-case mode extinction ratio degrades to 26 dB, the insertion loss climbs to 1.7 dB, and the MDL rises to 0.8 dB, respectively. Despite the introduction of core-dependent loss, the proposed device exhibits considerable performance stability. However, to enhance device performance, attention should be paid to power balance maintenance.

6.3.4 Summary

In conclusion, we have introduced a fiber-based all-optical unitary transformer suitable for implementation in MDM systems and various optical domains. Our proposed

device relies on multiple segments of MCFs, with phase controllers positioned before and after each segment. We have developed a straightforward yet effective algorithm to determine the ideal phase configuration for the phase shifters. Numerical simulations confirm the device's effectiveness and its suitability as a tunable mode MUX/DEMUX for few-mode fibers.

6.3.5 *The Proof that an Arbitrary Unitary Transformer Can Be Realized Based on the Proposed Structure*

In this part, a rigorous mathematical demonstration is presented to show that the proposed structure, detailed in the section, can achieve any arbitrary unitary transformation.

To give a solid proof, we must declare two theorems that have been thoroughly validated in reference [60] and establish a lemma.

Theorem I [60]: An arbitrary unitary $n \times n$ matrix (the group of $\mathbf{U}(n, \mathbb{C})$) can be realized by cascaded multiplication of the matrix $\mathbf{T}(i, j)$ (the group of $\mathbf{U}_{ij}$) and diagonal matrices (the group of $\mathbf{D}(n, \mathbb{C})$), where matrix $\mathbf{T}(i, j)$ is a unitary matrix in the form of

$$T(i, j) = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ & & \ddots & & & & \\ & & & \cos \theta & \cdots & \sin \theta & \\ & & & \vdots & 1 & \vdots & \\ & & & -\sin \theta & \cdots & \cos \theta & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \quad (6.46)$$

Theorem II [60]: $\mathbf{H}$ is a subgroup of $\mathbf{U}(n, \mathbb{C})$ which contains the diagonal matrix group $\mathbf{D}(n, \mathbb{C})$. If there is a matrix $\mathbf{A}$ in $\mathbf{H}$, with two of its elements $a_{ik} \neq 0$ and $a_{jk} \neq 0$ (with $i \neq j$) for some k , we have

$$\mathbf{H} \geq \mathbf{U}_{ij} \quad (6.47)$$

Lemma The transfer matrices of the circularly aligned MCF couplers and the transfer matrices of the phase shifter arrays, as well as their matrix product will form a subgroup of $\mathbf{U}(n, \mathbb{C})$.

Proof First of all, it is easy to see that all of the matrices within the set are unitary matrices. Also, it is straight forward to show that any diagonal matrices realized by the phase shifters have their inverse matrices to be contained within the set. Therefore,

to prove that the above matrix set is a subgroup of $\mathbf{U}(n, \mathbb{C})$, we need to prove that if we have one MCF coupler transfer matrix $\mathbf{A}$ in the set, we will have its inverse matrix $\mathbf{A}^{-1}$ to be contained in the set as well.

Afterwards, it can be easily proven that the transfer matrix $\mathbf{A}$ is a symmetrical matrix based on Eqs. (6.37), (6.38) and (6.39). We may choose a diagonal matrix $\mathbf{C}_1$ realized by the phase shifters such that

$$\mathbf{C}_1 = \begin{pmatrix} e^{j\theta_1} & & & \\ & e^{j\theta_2} & & \\ & & \ddots & \\ & & & e^{j\theta_n} \end{pmatrix}$$

$$\theta_m = -2 \times \text{angle}(A(1, m)) \quad (6.48)$$

where angle indicates the phase of a complex variable. It can be found that the matrix product $\mathbf{AC}_1\mathbf{A}$ has its first row and first column to be with only one non-zero element, i.e.

$$\mathbf{AC}_1\mathbf{A} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & E_{11} & & E_{1,n-1} \\ \vdots & & & \\ 0 & E_{n-1,1} & & E_{n-1,n-1} \end{pmatrix} \quad (6.49)$$

Since $\mathbf{A}$ and $\mathbf{C}_1$ are symmetrical and unitary, it can be concluded that the block matrix $\mathbf{E}$ is also symmetrical and unitary. Following the above steps, it is possible to find a diagonal matrix $\mathbf{C}_2$ which is realized by the phase shifters, such that

$$\begin{aligned} \mathbf{AC}_1\mathbf{AC}_2\mathbf{AC}_1\mathbf{A} &= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & E_{11} & & E_{1,n-1} \\ \vdots & & & \\ 0 & E_{n-1,1} & & E_{n-1,n-1} \end{pmatrix} \mathbf{C}_2 \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & E_{11} & & E_{1,n-1} \\ \vdots & & & \\ 0 & E_{n-1,1} & & E_{n-1,n-1} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & \cdots & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & F_{11} & & F_{1,n-2} \\ \vdots & & & & \\ 0 & 0 & F_{n-2,1} & & F_{n-2,n-2} \end{pmatrix} \end{aligned} \quad (6.50)$$

where the $(n-2) \times (n-2)$ block matrix $\mathbf{F}$ is still symmetrical and unitary. Hence, we may find the diagonal matrix $\mathbf{C}_3 \dots \mathbf{C}_n$ sequentially, such that

$$\mathbf{AC}_1\mathbf{A} \cdots \mathbf{AC}_{n-1}\mathbf{AC}_n\mathbf{AC}_{n-1}\mathbf{A} \cdots \mathbf{AC}_1\mathbf{A} = \mathbf{I} \quad (6.51)$$

where I is the identity matrix. Hence, we have

$$\mathbf{A}^{-1} = \mathbf{C}_1 \mathbf{A} \cdots \mathbf{A} \mathbf{C}_{n-1} \mathbf{A} \mathbf{C}_n \mathbf{A} \mathbf{C}_{n-1} \mathbf{A} \cdots \mathbf{A} \mathbf{C}_1 \mathbf{A} \quad (6.52)$$

i.e. $\mathbf{A}^{-1}$ can be realized by the matrix product of the transfer matrix $\mathbf{A}$ and the diagonal matrices $\mathbf{C}_1$ to $\mathbf{C}_n$. Hence, $\mathbf{A}^{-1}$ is contained in the set. Therefore, the set will form a subgroup of $U(n, C)$.

End of proof.

Based on the above two theorems and the lemma, it is possible to prove that our proposed structure can realize an arbitrary unitary transformer.

Proof From Theorem 1, since we have the tunable phase shifter arrays to realize the arbitrary diagonal matrices, it can be inferred that the key to realize an arbitrary unitary matrix is to realize an arbitrary matrix $\mathbf{T}(i,j)$ in the group of U_{ij} .

From Theorem 2, it can be inferred that to realize an arbitrary matrix $\mathbf{T}(i,j)$ in the group of U_{ij} , the key is to find a subgroup $\mathbf{H}$, which contains a matrix $\mathbf{A}$ with the non-zero elements of a_{ik} and a_{jk} .

Therefore, we propose to use the subgroup $\mathbf{H}$, which has been discussed in the lemma. Clearly, such a subgroup $\mathbf{H}$ contains the diagonal matrix group $D(n,C)$, and the transfer matrix $\mathbf{A}$ of the MCF coupler, which can be obtained by Eq. (6.39):

$$\mathbf{A}(z) = \mathbf{F} \mathbf{D}(z) \mathbf{F}^H$$

$$\mathbf{D}(z) = \begin{pmatrix} \exp(\lambda_1 z) & & & \\ & \exp(\lambda_2 z) & & \\ & & \ddots & \\ & & & \exp(\lambda_N z) \end{pmatrix} \quad (6.53)$$

where z is the length of the MCF coupler. As is discussed before, $\mathbf{A}$ is a symmetrical matrix.

From Eq. (6.53), we have

$$A_{ij}(z) = \sum_l F_{il} D_{ll}(z) F_{jl}^* \quad (6.54)$$

Clearly, $A_{ij}(z)$ is a non-zero function of z . Identifying the zeros of $A_{ij}(z)$ and do not choose those values for z , we have $A_{ij}(z) \neq 0$. Since $\mathbf{A}(z)$ is symmetrical, $A_{ji}(z) \neq 0$ if we have avoided the zeros of $A_{ij}(z)$. Such a matrix $\mathbf{A}(z)$ contains two non-zero elements, i.e. A_{ij} and A_{ji} . Thus, the proposed subgroup $\mathbf{H}$ contains matrix $\mathbf{T}(i,j)$.

We may further design a different MCF coupler, whose transfer matrix $\mathbf{M}$ contains the non-zero elements $M_{i'j'}$ and $M_{j'i'}$ to find the other matrix $\mathbf{T}(i',j')$ to be contained in the subgroup as well. In this way, an arbitrary matrix in the form of $\mathbf{T}(i,j)$ (with i, j being a integer from 1 to n and $i \neq j$) can be included in the subgroup $\mathbf{H}$. By Theorem 1, it is known that an arbitrary unitary matrix can be realized by the cascaded multiplication of the matrices within the subgroup.

Therefore, it is strictly proven that an arbitrary unitary transformer can be realized by the proposed structure given in the section.

End of proof.

From the discussion above, it can be inferred that the lengths of the MCF couplers should be chosen carefully to ensure more non-zero elements within the transfer matrix. If one wants to use identical MCF couplers, the length of the MCF couplers should be chosen so that no zero elements appear in the transfer matrix. Furthermore, the absolute values of the elements should be as close as possible.

We may choose the CL value to minimize the difference between the maximum absolute value of the elements and the minimum absolute value within the transfer matrix.

For example, to realize an arbitrary transformer in $U(2, C)$, we choose the coupling strength and the length product $CL = 0.25\pi$, so that we have the transfer matrix as

$$N_{opt}(2) = \begin{pmatrix} \frac{1}{\sqrt{2}} & j\frac{1}{\sqrt{2}} \\ j\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \quad (6.55)$$

In Eq. (6.50) all of the elements within the matrix have equal absolute values.

To realize an arbitrary transformer in $U(3, C)$, one may choose the optimal CL as $2/9\pi$, so that the transfer matrix have equal absolute values for the elements of the transfer matrix.

$$N_{opt}(3) = \frac{1}{\sqrt{3}} \begin{pmatrix} \exp(-j\frac{1}{18}\pi) & \exp(j\frac{11}{18}\pi) & \exp(j\frac{11}{18}\pi) \\ \exp(j\frac{11}{18}\pi) & \exp(-j\frac{1}{18}\pi) & \exp(j\frac{11}{18}\pi) \\ \exp(j\frac{11}{18}\pi) & \exp(j\frac{11}{18}\pi) & \exp(-j\frac{1}{18}\pi) \end{pmatrix} \quad (6.56)$$

To realize an arbitrary transformer in $U(4, C)$, one may choose the optimal CL as 0.25π , so that the transfer matrix have equal absolute values for the elements of the transfer matrix.

$$N_{opt}(4) = \begin{pmatrix} 0.5 & 0.5i & -0.5 & 0.5i \\ 0.5i & 0.5 & 0.5i & -0.5 \\ -0.5 & 0.5i & 0.5 & 0.5i \\ 0.5i & -0.5 & 0.5i & 0.5 \end{pmatrix} \quad (6.57)$$

6.3.6 Realization of a 3×3 Arbitrary Unitary Transformer

In this section, we discuss the possible number of $T(i, j)$ matrices (indicated by Eq. (6.46)) required to realize a 3×3 arbitrary unitary transformer. We will prove the following theorem to show that only three $T(i, j)$ matrices are required:

Theorem 3: An arbitrary unitary matrix $\mathbf{U}$ can be decomposed into the product of three unitary block matrices $\mathbf{A}^{(1)}$, $\mathbf{D}$, $\mathbf{A}^{(2)}$, i.e.

$$\begin{aligned}\mathbf{U} &= \mathbf{A}^{(1)} \mathbf{D} \mathbf{A}^{(2)} \\ \mathbf{A}^{(1)} &= \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} \\ & & a_{33}^{(1)} \end{pmatrix}, \mathbf{A}^{(2)} = \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} \\ & & a_{33}^{(2)} \end{pmatrix} \\ \mathbf{D} &= \begin{pmatrix} d_{11} & & \\ & d_{22} & d_{23} \\ & d_{32} & d_{33} \end{pmatrix}\end{aligned}\quad (6.58)$$

Proof Considering an arbitrary unitary matrix $\mathbf{U}$, we may find a block unitary matrix $\mathbf{C}$ such that their product results in a unitary matrix $\mathbf{V}$ whose element on the upper right corner V_{13} is zero, i.e.

$$\begin{aligned}\mathbf{U} &= \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix} \\ \mathbf{C} &= \begin{pmatrix} c_{11} & c_{12} & 0 \\ c_{21} & c_{22} & 0 \\ 0 & 0 & c_{33} \end{pmatrix} \\ c_{11} &= \frac{-u_{23}}{\sqrt{|u_{13}|^2 + |u_{23}|^2}}, c_{12} = \frac{u_{13}}{\sqrt{|u_{13}|^2 + |u_{23}|^2}}, \\ c_{21} &= -c_{12}^*, c_{22} = c_{11}^*, c_{33} = e^{j\eta} \\ \mathbf{V} &= \mathbf{CU} = \begin{pmatrix} v_{11} & v_{12} & 0 \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix}\end{aligned}\quad (6.59)$$

Since $\mathbf{V}$ is unitary, the rows of the matrix are orthogonal to each other, we have

$$\begin{aligned}v_{11}^* v_{21} + v_{12}^* v_{22} &= 0 \\ v_{11}^* v_{31} + v_{12}^* v_{32} &= 0\end{aligned}\quad (6.60)$$

With the matrix $\mathbf{V}$, it is possible to find a block matrix $\mathbf{B}$, whose elements fulfill the following:

$$\mathbf{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ & & b_{33} \end{pmatrix}$$

$$\begin{aligned}
 b_{11} &= \frac{v_{11}^*}{\sqrt{|v_{11}|^2 + |v_{12}|^2}}, b_{21} = \frac{v_{12}^*}{\sqrt{|v_{11}|^2 + |v_{12}|^2}}, \\
 b_{12} &= -b_{21}^*, b_{22} = b_{11}^*, b_{33} = e^{j\zeta}
 \end{aligned} \tag{6.61}$$

The product of $\mathbf{VB}$ will result in a matrix $\mathbf{D}$, which has two of its elements on the bottom left corner to be zero:

$$\begin{aligned}
 \mathbf{D} = \mathbf{VB} &= \begin{pmatrix} v_{11} & v_{12} & 0 \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ & b_{33} \end{pmatrix} \\
 &= \begin{pmatrix} d_{11} & d_{12} & d_{13} \\ 0 & d_{22} & d_{23} \\ 0 & d_{32} & d_{33} \end{pmatrix}
 \end{aligned} \tag{6.62}$$

Since $\mathbf{D}$ is unitary, we have

$$\begin{aligned}
 |d_{11}| &= 1 \\
 d_{12} &= 0 \\
 d_{13} &= 0
 \end{aligned} \tag{6.63}$$

Hence, we have matrix $\mathbf{U}$ to be decomposed into:

$$\begin{aligned}
 \mathbf{U} &= \mathbf{C}^{-1} \mathbf{D} \mathbf{B}^{-1} \\
 &= \begin{pmatrix} c_{11}^* & c_{21}^* & 0 \\ c_{12}^* & c_{22}^* & 0 \\ 0 & 0 & c_{33}^* \end{pmatrix} \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & d_{23} \\ 0 & d_{32} & d_{33} \end{pmatrix} \begin{pmatrix} b_{11}^* & b_{21}^* \\ b_{12}^* & b_{22}^* \\ & b_{33}^* \end{pmatrix}
 \end{aligned} \tag{6.64}$$

Clearly, Eq. (6.64) is in accordance with Eq. (6.58).

End of proof.

Combining the lemma and the Theorems 1, 2 and 3, and using the procedures indicated in [60], it is possible to determine the maximum number of MCF couplers and phase-shifters required to realize an arbitrary 3X3 unitary transform.

6.4 Orbital Angular Momentum Mode Sorting Based on a Hybrid Radial-Angular Hybrid Lens

Part of the section reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61].

The previous sections discussed the mode multiplexers/demultiplexer designed for the different shape of waveguides. In this section, we introduce a mode multiplexer/demultiplexer tailored for orbital angular momentum (OAM) modes. OAM modes exhibit a phase distribution of the form $\exp(jl\theta)$, mirroring the characteristics of a plane wave in Cartesian coordinates. Similar to a conventional lens that focuses a plane wave onto a focal plane, an angular lens can concentrate an OAM beam in the angular domain. However, it often results in a protracted tail due to the inadequate angular focal condition for beams that aren't ring-shaped. Here, we propose a hybrid lens operating in both the angular and radial domains. This radial lens, featuring a specially designed radially distributed phase, ensures that the angular focal condition is met for beams with varying beam waists or radial field distributions. This significantly enhances the sorting efficiency of OAM modes by the angular lens. Efficient discrimination among different OAM modes becomes feasible using this single optical component—our proposed hybrid radial-angular lens—with an OAM mode inter-mode crosstalk of just 3.7% when the topological charge difference is 3.

6.4.1 Introduction

Orbital angular momentum multiplexing (OAMM) has emerged as a leading technique for boosting the capacity of optical communication systems [19]. Besides its applications in optical communications [19], it finds utility in quantum information [62] and optical signal processing [63, 64].

While there are various methods for generating OAM modes, such as spiral phase plates (SPPs) [21], circularly arranged cores for coherent combining [25], ring resonators [39], conversion from spin angular momentum to OAM [28], and index matching of OAM modes to fiber modes [27], the challenge lies in OAM mode sorting. This is because it enables intensity detection, which remains the prevalent detection method unless costly wave-front sensors are employed. Many approaches have been suggested for OAM mode sorting, including interference methods [22, 65], diffraction techniques based on SPPs/holograms or specific apertures [62, 66, 67], and conformal transforms [68, 69]. Each method has its strengths and weaknesses. For instance, the commonly used SPP/hologram approach offers high conversion/detection efficiency for OAM modes, but each OAM mode requires a unique SPP, making high-volume multiplexing costly and complex. The conformal transform method needs two elements: one for log-polar transformation and the other for phase correction. Recently, an angular lens-based method has been introduced for OAM mode sorting [70, 71]. This innovative approach employs a lens with a parabola index distribution in the angular domain, focusing OAM modes on the “angular focal plane.” However, a significant limitation of the angular lens is its suitability for sorting OAM modes with a specific radius [71], its performance and efficiency decline when the beam radius doesn't match the designed target, especially for non-ring-shaped beams.

In this study, we introduce a new method for distinguishing OAM modes. The concept involves the integration of an angular lens with a radial lens. This combination allows for the focusing of beams carrying OAM topological charges in the angular domain, overcoming the limitations of previously reported standalone angular lenses [70–72]. The radial lens enables converging the beam to a precise radius on the output plane, which is inversely related to its radius on the input plane. This setup ensures that waves of varying radii on both input and output planes meet the focusing criteria of the angular lens. By implementing this hybrid lens design, sorting efficiency and overall performance are notably enhanced. Furthermore, this hybrid lens facilitates single-component-based OAM mode sorting. It's important to clarify that our hybrid lens differs from the spin–orbit transformation meta-lens described in [73] and the multi-focal hybrid lens in [74], as it features a single focal length and doesn't include spin–orbit conversion.

The radial lens enables converging the beam to a precise radius on the output plane, which is inversely related to its radius on the input plane. This setup ensures that waves of varying radii on both input and output planes meet the focusing criteria of the angular lens. By implementing this hybrid lens design, sorting efficiency and overall performance are notably enhanced. Furthermore, this hybrid lens facilitates single-component-based OAM mode sorting. It is important to clarify that our hybrid lens differs from the spin–orbit transformation meta-lens described in [73] and the multi-focal hybrid lens in [74], as it features a single focal length and does not include spin–orbit conversion.

6.4.2 Theory

We consider the scale wave propagation in free space, which is formulated as:

$$E(\rho, \varphi, z) = \frac{\exp(jkz)}{j\lambda z} e^{j\frac{k}{2z}\rho^2} \int_0^{+\infty} \int_0^{2\pi} E(\rho_1, \varphi_1, 0) T(\rho_1, \varphi_1) e^{j\frac{k}{2z}\rho_1^2} e^{-j\frac{k\rho_1\rho}{z} \cos(\varphi - \varphi_1)} \rho_1 d\rho_1 d\varphi_1, \quad (6.65)$$

where $\rho_1, \varphi_1, \rho, \varphi$, stand for the polar coordinates on the input and output planes, z denotes the propagation distance, k denotes the wave number, λ denotes the wavelength, and T denotes the phase introduced by the lens.

For an angular lens, its phase should follow:

$$T(\rho_1, \varphi_1) = \exp\left(-j\frac{N}{4\pi}\varphi_1^2\right) \exp(-j\beta\rho_1), \quad (6.66)$$

where N denotes a scaling factor and β denotes a radial correction factor for the angular lens [70]. It is anticipated that the OAM modes will exhibit azimuthal convergence. As is demonstrated in Sect. 6.6.5 (disregarding the correction factor in the analysis) that a ring-shape beam with radius r_1 on the input plane will concentrate in a circular sector with radius r_2 on the output plane at the focal length f ,

$$f = \frac{2\pi k r_1 r_2}{N}. \quad (6.67)$$

The angle of the circular segment is influenced by the topological charge; specifically, the output wave converges within a segment of radius r_2 and as the topological charge varies.

Assuming a ring-shaped beam of radius a is at the input and an output ring of the same radius, we have $f = \frac{2\pi k a^2}{N}$. This aligns with Eq. (6.70) in [71]. The comprehensive derivation of Eq. (6.67) is presented in Sect. 6.6.5.

However, Eq. (6.67) is applicable only to ring-shaped beams on both input and output planes. In practical situations, beams often deviate from a perfect ring shape, and various rings on the input and output planes possess differing angular focal lengths. Consequently, the performance deteriorates significantly, leading to extended tails after the OAM modes are focused [70, 71].

To address this, we propose integrating a radial ring-lens [75] with the angular lens. This combination aims to confine the output beam to a ring-shaped area. If the phase of lens forms a radial lens, i.e.,

$$T(\rho_1, \varphi_1) = \exp\left(-j \frac{k}{2f} (\rho_1 - a)^2\right). \quad (6.68)$$

At the focal length f , the beam $E(\rho, \varphi, f)$ concentrates into a ring with the radius of a . By combining both radial and angular lenses, we achieve superior OAM angular focusing compared to using a pure angular lens. This is because the hybrid approach ensures that the output beam maintains a distinct ring radius.

Figure 6.21 presents the simulation results comparing the angular lens and the basic hybrid radial-angular lens. Evidently, when the input beam is Gaussian-shaped, the output beam produced by a pure angular lens exhibits a long tail. The introduction of the hybrid radial-angular lens mitigates this tail. However, the output beam still displays some side lobes, primarily because the input beam lacks a perfect ring shape. Equation (6.67) suggests that waves with varying radii at the input are likely to have differing focal lengths, explaining this phenomenon.

To enhance performance further, we propose to adjust the focal radius of the ring a based on the input plane's radius. This ensures that Eq. (6.67) is applicable to the waves with varying radii on the input plane, as given by

$$a(\rho_1) = \frac{r_0^2}{\rho_1}. \quad (6.69)$$

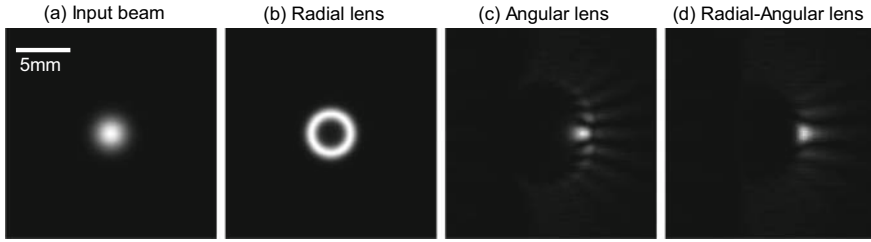


Fig. 6.21 Simulation of the mode sorter with the simple hybrid lens with the angular lens and the radial direction ring lens. **a** The input beam is the Gaussian shape beam, **b** the focus beam with a radial lens, **c** the focused beam with a pure angular lens, **d** the focused beam with the simple hybrid radial-angular lens. The parameters are as follows. The wavelength of the beam is 671 nm, the beam waist w_0 is 1.9 mm, the parameter N is 50, a is 1.9 mm, and β is 2/mm. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

With this approach, waves on the input beam with a radius of ρ_1 will converge onto a ring with a corresponding radius $\frac{r_0^2}{\rho_1}$ on the output plane. Consequently, waves at different radii are expected to share the same focal length $f = \frac{2\pi ka(\rho_1)\rho_1}{N} = \frac{2\pi kr_0^2}{N}$.

The enhanced hybrid radial-angular lens operates based on the principle illustrated in Fig. 6.22. By linking the input radius to the output focal radius, we can ensure that the product of wave radii on the input and output planes remains constant, specifically r_0^2 . This adaptation enables Eq. (6.67) to be applicable to non-ring-shaped OAM beams.

In conclusion, hybrid lens has the following phase distribution

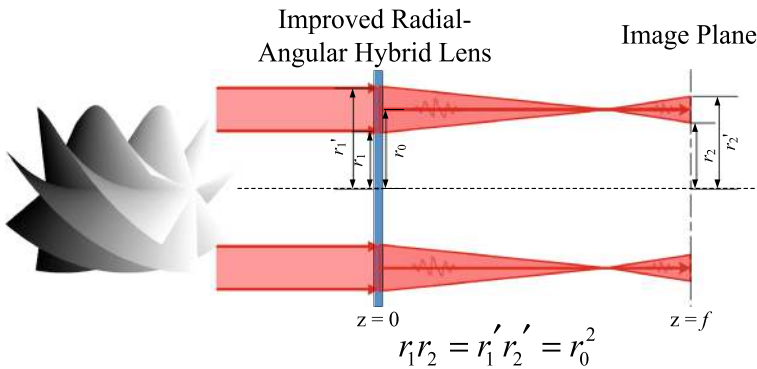


Fig. 6.22 Illustration of the focal condition for the improved hybrid radial-angular lens. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

$$T(\rho_1, \varphi_1) = \exp\left(-j\frac{k}{2f}\left(\rho_1 - \frac{r_0^2}{\rho_1}\right)^2\right) \exp\left(-j\frac{N}{4\pi}\varphi_1^2\right) \exp(-j\beta\rho_1),$$

$$f = \frac{2\pi kr_0^2}{N}.$$
(6.70)

6.4.3 Results and Discussions

The proposed OAM mode sorter based on the hybrid radial-angular lens has been realized experimentally, with the setup demonstrated in Fig. 6.23.

A continuous wave laser, operating at a wavelength of 671 nm, serves as the light source. This laser beam undergoes magnification via a laser beam expander and has its intensity adjusted by an attenuator. The beam then illuminates a SLM with the pixel size of $9.4 \mu\text{m} \times 9.4 \mu\text{m}$, and the grid of 800×600 . After acquiring a helical wavefront from the first SLM, the beam transforms into an OAM mode. The second SLM has the phase according to Eq. (6.70), which is depicted in Fig. 6.24. It functions as our improved hybrid radial-angular lens. Its parameters are set as follows: $N = 50$, $r_0 = 0.6 \text{ mm}$, and $\beta = 2/\text{mm}$.

This phase mask enables OAM mode sorting, which is demonstrated on the output images captured by a CMOS camera (grid resolution of 2160×4096 , grid size of $3.45 \mu\text{m} \times 3.45 \mu\text{m}$), as shown in Fig. 6.25. As is anticipated, the output beam focuses

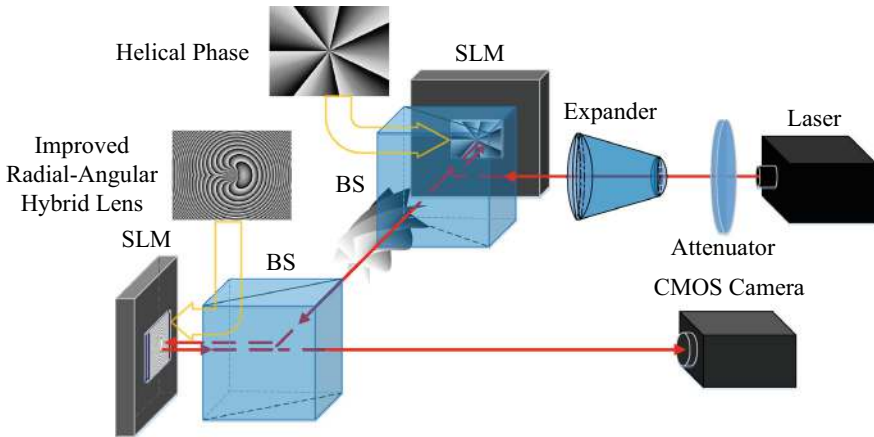


Fig. 6.23 Experimental setup. The 671 nm laser has its beam expanded by an expander. A helical wavefront is added to the first SLM, which converts the beam into an OAM mode. The OAM mode is sorted by the second SLM which has the phase of the improved radial-angular lens on it. SLM: spatial light modulator. BS: beam splitter. CMOS camera: Complementary Metal Oxide Semiconductor camera. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

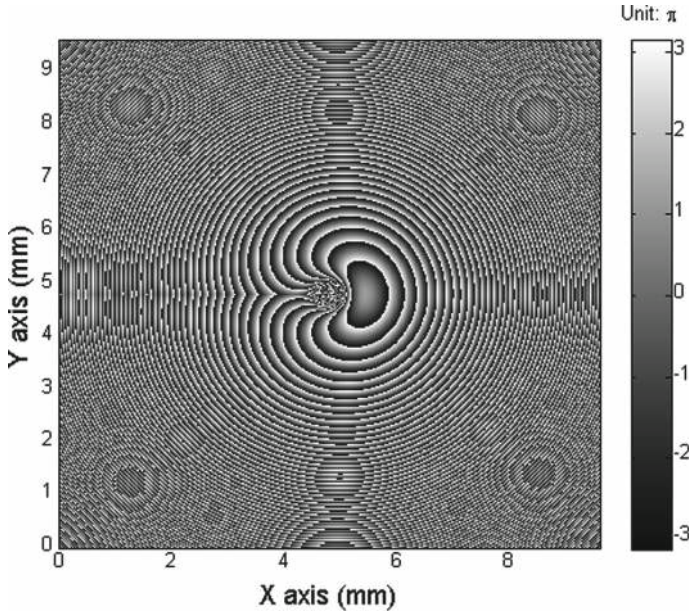


Fig. 6.24 Illustration of the phase distribution of the improved radial-angular hybrid lens. The parameters are as follows. The wavelength of the beam is 671 nm, the parameter N is 50, r_0 is 0.6 mm, and β is 2/mm. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

on a sector centered at 0 degrees when the input beam carries a topological charge of 0. As the topological charge increases, the output beam rotates counterclockwise, and vice versa. Compared to mode sorting using a pure angular lens in Fig. 6.21 and Ref. [70, 71], our sorted beam maintains a pencil-like shape without an extended tail in the azimuthal direction. Since OAM modes with varying topological charges occupy distinct sectors, mode sorting can be achieved through targeted measurements within specific sectors, which is similar to [68] but differs from more complex algorithmic approaches in [76, 77]. Thus, our enhanced hybrid radial-angular lens proves effective as an OAM mode sorter. When a beam consists of two OAM modes with different topological charges, our sorter can separate them in the azimuthal domain. In our experiment, we generated a combined OAM mode with topological charges of -2 and $+2$ using the first SLM, whose pixels were randomly filled with two phase patterns at equal probability [78]. The bottom of Fig. 6.25 demonstrates that this combined ± 2 OAM mode beam is effectively split into two pencil-like beams with distinct rotation angles. If a fractional-order OAM mode is used as input, multiple pencil-like beams appear on the output plane. This is because fractional-order OAM modes can be decomposed into a series of integer-order OAM modes, as is noted in [79, 80].

Figure 6.26 illustrates the crosstalk matrix to quantitatively assess the performance of the proposed OAM sorter. This evaluation involves dividing the detector plane into

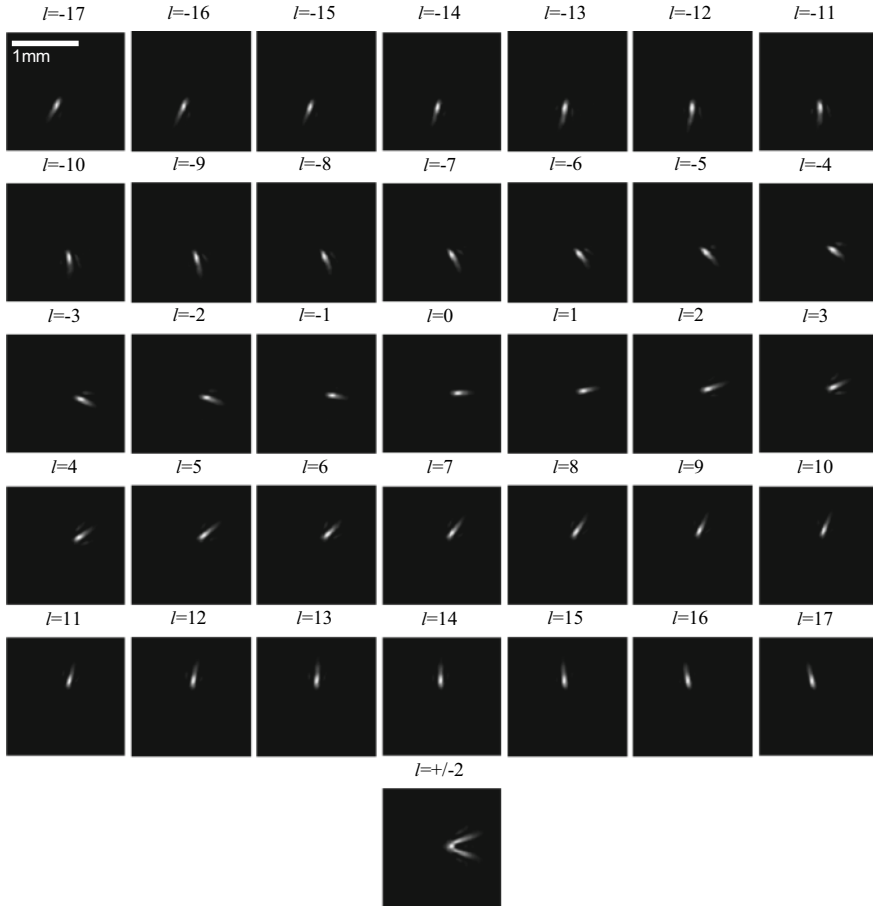


Fig. 6.25 Mode sorting with the proposed hybrid lens for the different OAM modes. The images are captured by the CMOS camera. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

distinct sectors, each linked to a specific OAM mode output. The ratio of the sorted mode's power within its designated sector determines the mode's correct detection probability. Due to mode overlap at the output, the average accuracy for detecting each mode stands at 49%. Additionally, there's an average crosstalk of 21% between adjacent OAM modes. Remarkably, when the topological charge separation of the OAM modes is 3, as is demonstrated in Fig. 6.27. This remarkable performance surpasses the mode sorting abilities of the pure angular-lens-based OAM mode sorter [70], which exhibits an average crosstalk of 16.5% even with a topological charge separation of 4 for the OAM modes.

To precisely sort the OAM modes without topological charge separation, i.e., the OAM mode ...−2, −1, 0, 1, 2, ... one may implement the OAM mode multiplier

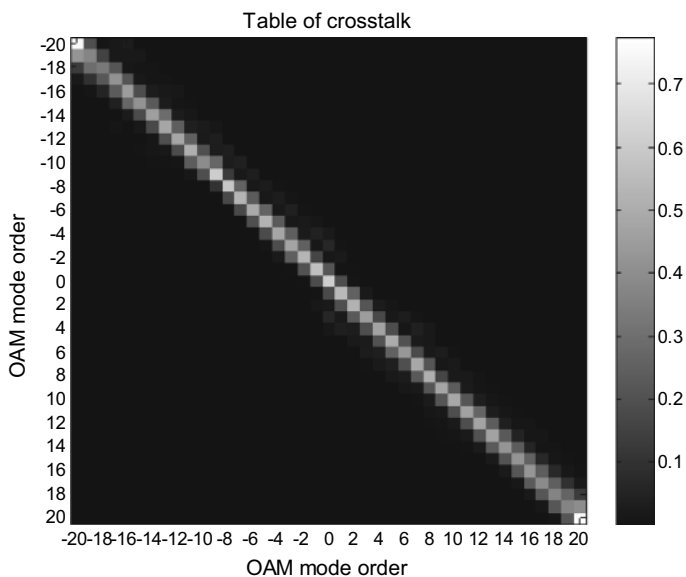


Fig. 6.26 The table of crosstalk for different OAM modes. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

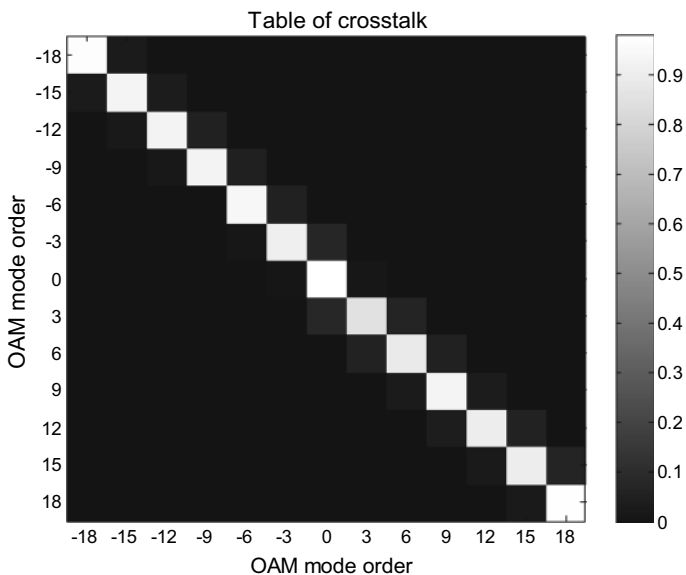


Fig. 6.27 The table of crosstalk for different OAM modes with the topological charge separation of 3. This figure is reprinted/adapted with permission from [Opt. Express 30, 9703–9713 (2022)] © The Optical Society of America [61]

[64]. This multiplier can transform the topological charge l to lM . If $M = 3$, an exceptionally high correct detection probability of 92.5% and a notably low crosstalk of 3.7% can be attained.

6.4.4 Summary

In conclusion, we have introduced a novel hybrid lens capable of focusing OAM modes in the angular direction. This hybrid lens combines an angular lens, featuring a quadratic phase in the azimuth domain, with a radial lens whose focal radius varies inversely with the input wave radius. The hybrid lens performs quite efficiently in OAM mode sorting, producing a pencil-like output beam whose azimuth angle rotates counterclockwise as the topological charge rises. Remarkably, when the mode topological charge separation is 3, crosstalk between adjacent modes remains at just 4%. Compared to reported OAM mode sorters, our approach simplifies the sorting process by requiring only one optical element, i.e., the proposed hybrid radial-angular lens, thereby simplifying the OAM mode sorting system and cutting costs. Performance can be further enhanced by inserting an OAM multiplier before the hybrid lens. This innovative lens can be implemented via an SLM, as is demonstrated in our study, or through alternative methods like holograms or meta-lenses.

6.4.5 The Detailed Derivation for the Angular Lens Equation with the Specific Input Ring and Output Ring Radii

In this section, the detailed derivation for the angular lens equation with the specific input ring and output ring radii (i.e., Eq. (6.67) in the main text) is provided. In addition, the radial phase distribution for the improved angular-radial lens is derived (i.e., Eq. (6.70) in the main text).

The scalar beam propagation in free space can be characterized by Eq. (6.65) (rewritten as Eq. (6.71))

$$E(\rho, \varphi, z) = \frac{\exp(jkz)}{j\lambda z} e^{j\frac{k}{2z}\rho^2} \int_0^{+\infty} \int_0^{2\pi} E(\rho_1, \varphi_1, 0) T(\rho_1, \varphi_1) e^{j\frac{k}{2z}\rho_1^2} e^{-j\frac{k\rho_1\rho}{z} \cos(\varphi - \varphi_1)} \rho_1 d\rho_1 d\varphi_1, \quad (6.71)$$

where $\rho_1, \varphi_1, \rho, \varphi$, are the polar coordinates on the input and output planes, z the propagation distance, k the wave number, λ the wavelength, and T the phase introduced by the lens and is assumed as the radially quadratic phase

$$T(\varphi_1) = \exp\left(-\frac{N}{4\pi}\varphi_1^2\right), \quad (6.72)$$

where N is a scaling factor. Without loss of generality, we assume N to be an even integer. Since,

$$\exp(-jz \cos \varphi) = \sum_{n=-\infty}^{+\infty} (-j)^n e^{jn\varphi} J_n(z), \quad (6.73)$$

and the approximated formula for Bessel function when $z < 1$

$$\begin{aligned} J_n(z) &\approx \sqrt{\frac{2}{\pi z}} \left(\cos\left(z - \frac{2n+1}{4}\pi\right) - \sin\left(z - \frac{2n+1}{4}\pi\right) \frac{(4n^2-1)}{4} \frac{1}{2z} \right) \\ &= \sqrt{\frac{2}{\pi z}} \operatorname{Re} \left(e^{j(z - \frac{2n+1}{4}\pi)} \left(1 + j \frac{(4n^2-1)}{4} \frac{1}{2z} \right) \right) \\ &\approx \frac{1}{2} \sqrt{\frac{2}{\pi z}} \left(e^{j(z - \frac{2n+1}{4}\pi)} e^{j \frac{(4n^2-1)}{4} \frac{1}{2z}} + e^{-j(z - \frac{2n+1}{4}\pi)} e^{-j \frac{(4n^2-1)}{4} \frac{1}{2z}} \right), \end{aligned} \quad (6.74)$$

we have

$$\begin{aligned} E(\rho, \varphi, z) &\approx \frac{\exp(jkz)}{j\lambda z} \int_0^{+\infty} \int_0^{2\pi} E(\rho_1, \varphi_1, 0) \exp\left(-\frac{N}{4\pi}\varphi_1^2\right) \sqrt{\frac{1}{2\pi \frac{k\rho_1\rho}{z}}} \\ &\times \sum_{n=-\infty}^{+\infty} e^{jn(\varphi-\varphi_1)} \left((-1)^n e^{-j\frac{\pi}{4}} e^{j\frac{k(\rho_1+\rho)^2}{2z}} e^{j\frac{(4n^2-1)}{4} \frac{z}{2k\rho_1\rho}} + e^{j\frac{\pi}{4}} e^{j\frac{k(\rho_1-\rho)^2}{z}} e^{-j\frac{(4n^2-1)}{4} \frac{z}{2k\rho_1\rho}} \right) \\ \rho_1 d\rho_1 d\varphi_1 &\approx \frac{\exp(jkz)}{j\lambda z} \int_0^{+\infty} \int_0^{2\pi} E(\rho_1, \varphi_1, 0) \exp\left(-\frac{N}{4\pi}\varphi_1^2\right) \\ &\sqrt{\frac{1}{2\pi \frac{k\rho_1\rho}{z}}} \sum_{n=-\infty}^{+\infty} e^{jn(\varphi-\varphi_1)} e^{j\frac{\pi}{4}} e^{j\frac{k(\rho_1-\rho)^2}{z}} e^{-j\frac{(4n^2-1)}{4} \frac{z}{2k\rho_1\rho}} \rho_1 d\rho_1 d\varphi_1. \end{aligned} \quad (6.75)$$

The approximation is based on the fact that $e^{j\frac{k(\rho_1+\rho)^2}{2z}}$ oscillates much faster than $e^{j\frac{k(\rho_1-\rho)^2}{z}}$ when ρ_1 and ρ are comparable. Therefore, the former can be neglected in the integral.

Considering the contribution of the input wave with the radius of r_1 and the wave at the output plane with the radius of r_2 , we have

$$E(r_2, \varphi, z) \propto \int_0^{2\pi} \phi(\varphi_1) T(\varphi_1) \sum_{n=-\infty}^{+\infty} (e^{jn(\varphi-\varphi_1)}) e^{-j \frac{(4n^2-1)}{4} \frac{z}{2kr_1r_2}} d\varphi_1, \quad (6.76)$$

where $\phi(\varphi_1)$ is the azimuth distribution function for $E(\rho_1, \varphi_1, 0)$. Let $\frac{z}{2kr_1r_2} = \frac{\pi}{N}$, and $n = gN + m$ ($0 = < m < N$), we have

$$\begin{aligned} & \sum_{n=-\infty}^{+\infty} (e^{jn(\varphi-\varphi_1)}) e^{-j \frac{\pi n^2}{N}} \\ &= \sum_{g=-\infty}^{+\infty} \sum_{m=0}^{N-1} (e^{j(m+gN)(\varphi-\varphi_1)}) e^{-j \frac{\pi (m+gN)^2}{N}} \\ &= \sum_{m=0}^{N-1} (e^{jm(\varphi-\varphi_1)} e^{-j \frac{\pi m^2}{N}}) \sum_{g=-\infty}^{+\infty} e^{jgN(\varphi-\varphi_1)} \\ &= \sum_{m=0}^{N-1} (e^{jm(\varphi-\varphi_1)} e^{-j \frac{\pi m^2}{N}}) \frac{2\pi}{N} \sum_{g=-\infty}^{+\infty} \delta\left(\varphi - \varphi_1 - g \frac{2\pi}{N}\right) \\ &= \frac{2\pi}{N} \sum_{m=0}^{N-1} (e^{jmg \frac{2\pi}{N}} e^{-j \frac{\pi m^2}{N}}) \sum_{g=-\infty}^{+\infty} \delta\left(\varphi - \varphi_1 - g \frac{2\pi}{N}\right). \end{aligned} \quad (6.77)$$

Based on Eq. (6.77), we have the integral in the azimuth domain as

$$\begin{aligned} E(r_2, \varphi, z) \propto I &= \int_0^{2\pi} \phi(\varphi_1) T(\varphi_1) \sum_{n=-\infty}^{+\infty} (e^{jn(\varphi-\varphi_1)}) e^{-j \frac{\pi n^2}{N}} d\varphi_1 \\ &= \frac{2\pi}{N} \int_0^{2\pi} \phi(\varphi_1) \exp\left(-\frac{N}{4\pi} \varphi_1^2\right) \\ & \quad \sum_{m=0}^{N-1} (e^{jm(\varphi-\varphi_1)} e^{-j \frac{\pi m^2}{N}}) \sum_{g=-\infty}^{+\infty} \delta\left(\varphi - \varphi_1 - g \frac{2\pi}{N}\right) d\varphi_1 \\ &= \frac{2\pi}{N} \sum_{g=-\infty}^{+\infty} \phi\left(\varphi - g \frac{2\pi}{N}\right) \exp\left(-\frac{N}{4\pi} \left(\varphi - g \frac{2\pi}{N}\right)^2\right) \sum_{m=0}^{N-1} (e^{jmg \frac{2\pi}{N}} e^{-j \frac{\pi m^2}{N}}). \end{aligned} \quad (6.78)$$

Let $\varphi = \frac{2\pi l}{N}$, $l = 0 \dots N-1$, which means samples are taken evenly on the azimuth direction on the output plane, and let $g = g'N + m'$ ($0 = < m' < N$), while

considering the fact that $\varphi = \varphi \pm 2\pi g'$, we have

$$\begin{aligned}
 I\left(\varphi = \frac{2\pi l}{N}\right) &= \frac{2\pi}{N} \sum_{g'=-\infty}^{+\infty} \sum_{m'=0}^{N-1} \phi\left(\frac{2\pi}{N}(l-m')\right) \\
 &\exp\left(-\frac{N}{4\pi}\left(\frac{2\pi}{N}(l-m')\right)^2\right) \sum_{m=0}^{N-1} \left(e^{jmm'\frac{2\pi}{N}} e^{-j\frac{\pi m^2}{N}}\right) \\
 &\propto \sum_{m'=0}^{N-1} \phi\left((l-m')\frac{2\pi}{N}\right) \exp\left(-(l-m')^2\frac{\pi}{N}\right) \sum_{m=0}^{N-1} \left(e^{jmm'\frac{2\pi}{N}} e^{-j\frac{\pi m^2}{N}}\right). \quad (6.79)
 \end{aligned}$$

Let $k = l - m'$, we have

$$\begin{aligned}
 I\left(\varphi = \frac{2\pi l}{N}\right) &\propto \sum_{k=0}^{N-1} \phi\left(\frac{2\pi k}{N}\right) \exp\left(-\frac{k^2\pi}{N}\right) \sum_{m=0}^{N-1} \left(e^{jm(l-k)\frac{2\pi}{N}} e^{-j\frac{\pi m^2}{N}}\right) \\
 &= \frac{2\pi}{N} e^{-j\frac{\pi l^2}{N}} \sum_{k=0}^{N-1} \phi\left(\frac{2\pi k}{N}\right) e^{-j\frac{2\pi kl}{N}} \sum_{m=0}^{N-1} e^{-j\frac{\pi(k+m-l)^2}{N}} \\
 &= \frac{2\pi}{\sqrt{N}} e^{-j\frac{\pi}{4}} e^{-j\frac{\pi l^2}{N}} \sum_{k=0}^{N-1} \phi\left(\frac{2\pi k}{N}\right) e^{-\frac{2\pi kl}{N}}. \quad (6.80)
 \end{aligned}$$

Equation (6.80) suggests that the two terms constitute the discrete Fourier transform (DFT) pairs, i.e.,

$$I\left(\varphi = \frac{2\pi l}{N}\right) = \frac{2\pi}{\sqrt{N}} e^{-j\frac{\pi}{4}} e^{-j\frac{\pi l^2}{N}} DFT\left(\phi\left(\frac{2\pi k}{N}\right)\right). \quad (6.81)$$

So, with the change of the topological charge, one may expect the field concentration on the different angles due to the property of N point DFT. In order to guarantee the interference for all r_1 and r_2 at the distance z , one must ensure

$$r_1 r_2 = \frac{Nz}{2k\pi}, \quad (6.82)$$

which is Eq. (6.67) in the main text. To make sure that the wave with the radius of r_1 on the input plane goes to the radius of r_2 on the output plane, the radial lens is introduced at r_1 with the target output radius as

$$r_2(r_1) = \frac{r_0^2}{r_1}, \quad (6.83)$$

with

$$r_0^2 = \frac{Nz}{2k\pi}, \quad (6.84)$$

which is Eq. (6.70).

References

1. J. Zhou, P. Gallion, A novel mode multiplexer/demultiplexer for multi-core fibers. *IEEE Photonics Technol. Lett.* **25**(13), 1214–1217 (2013)
2. H.R. Stuart, Dispersive multiplexing in multimode optical fiber. *Science* **289**(5477), 281–283 (2000)
3. B.C. Thomsen, MIMO enabled 40 Gb/s transmission using mode division multiplexing in multimode fiber, in *2010 Conference on Optical Fiber Communication, Collocated National Fiber Optic Engineers Conference (OFC/NFOEC)* (2010)
4. R. Ryf, A. Sierra, R.J. Essiambre, A.H. Gnauck, S. Randel, M. Esmaelpour, S. Mumtaz, P.J. Winzer, R. Delbue, P. Pupalakis, A. Sureka, T. Hayashi, T. Taru, T. Sasaki, Coherent 1200-km 6×6 MIMO mode-multiplexed transmission over 3-core microstructured fiber, in *2011 37th European Conference and Exhibition on Optical Communications (ECOC 2011)* (2011)
5. S. Randel, M. Magarini, R. Ryf, R.J. Essiambre, A.H. Gnauck, P.J. Winzer, T. Hayashi, T. Taru, T. Sasaki, MIMO-based signal processing of spatially multiplexed 112-Gb/s PDM-QPSK signals using strongly-coupled 3-core fiber, in *2011 37th European Conference and Exhibition on Optical Communications (ECOC 2011)* (2011)
6. F. Saitoh, K. Saitoh, M. Koshiba, A design method of a fiber-based mode multi/demultiplexer for mode-division multiplexing. *Opt. Express* **18**(5), 4709–4716 (2010)
7. Y. Kokubun, and M. Koshiba, Novel multi-core fibers for mode division multiplexing: proposal and design principle. *IEICE Electron. Express* **6**(8), 522–528 (2009)
8. E. Kapon, J. Katz, A. Yariv, Supermode analysis of phase-locked arrays of semiconductor-lasers. *Opt. Lett.* **9**(4), 125–127 (1984)
9. L.B. Soldano, E.C.M. Pennings, Optical multimode interference devices based on self-imaging—principles and applications. *J. Lightwave Technol.* **13**(4), 615–627 (1995)
10. J.M. Heaton, R.M. Jenkins, General matrix theory of self-imaging in multimode interference (MMI) couplers. *IEEE Photonics Technol. Lett.* **11**(2), 212–214 (1999)
11. A.R. Gupta, K. Tsutsumi, J. Nakayama, Synthesis of Hadamard transformers by use of multimode interference optical waveguides. *Appl. Opt.* **42**(15), 2730–2738 (2003)
12. S.Y. Tseng, Y. Kim, C.J.K. Richardson, J. Goldhar, Implementation of discrete unitary transformations by multimode waveguide holograms. *Appl. Opt.* **45**(20), 4864–4872 (2006)
13. J.B. Park, D.M. Yeo, S.Y. Shin, Variable optical mode generator in a multimode waveguide. *IEEE Photonics Technol. Lett.* **18**(17–20), 2084–2086 (2006)
14. A.L.Y. Low, Y.S. Yong, A.H. You, S.F. Chien, C.F. Teo, A five-order mode converter for multimode waveguide. *IEEE Photonics Technol. Lett.* **16**(7), 1673–1675 (2004)
15. E. Simova, I. Golub, Polarization splitter/combiner in high index contrast Bragg reflector waveguides. *Opt. Express* **11**(25), 3425–3430 (2003)
16. X. Yang, Y. Yang, Q. Cai, Y. Zhao, A. Fang, W. Hu, A. Xu, Primary experiments on 2-D and 1-D fiber-type optical phased array
17. J.E. Antonio-Lopez, A. Castillo-Guzman, D.A. May-Arrioja, R. Selvas-Aguilar, P. LikamWa, Tunable multimode-interference bandpass fiber filter. *Opt. Lett.* **35**(3), 324–326 (2010)
18. J.H. Zhou, OAM states generation/detection based on the multimode interference effect in a ring core fiber. *Opt. Express* **23**(8), 10247–10258 (2015)
19. J. Wang, J.Y. Yang, I.M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y.X. Ren, Y. Yue, S. Dolinar, M. Tur, A.E. Willner, Terabit free-space data transmission employing orbital angular momentum multiplexing. *Nat. Photonics* **6**(7), 488–496 (2012)

20. C. Paterson, Atmospheric turbulence and orbital angular momentum of single photons for optical communication. *Phys. Rev. Lett.* **94**(15) (2005)
21. M.W. Beijersbergen, R.P.C. Coerwinkel, M. Kristensen, J.P. Woerdman, Helical-wavefront laser beams produced with a spiral phaseplate. *Opt. Commun.* **112**(5), 321–327 (1994)
22. J. Leach, M.J. Padgett, S.M. Barnett, S. Franke-Arnold, J. Courtial, Measuring the orbital angular momentum of a single photon. *Phys. Rev. Lett.* **88**(25) (2002)
23. M.J. Strain, X.L. Cai, J.W. Wang, J.B. Zhu, D.B. Phillips, L.F. Chen, M. Lopez-Garcia, J.L. O'Brien, M.G. Thompson, M. Sorel, S.Y. Yu, Fast electrical switching of orbital angular momentum modes using ultra-compact integrated vortex emitters. *Nat. Commun.* **5** (2014)
24. B.B. Guan, R.P. Scott, N.K. Fontaine, T.H. Su, C. Ferrari, M. Cappuzzo, F. Klemens, B. Keller, M. Earnshaw, S.J.B. Yoo, Integrated optical orbital angular momentum multiplexing device using 3-D waveguides and a silica PLC, in *2013 Conference on Lasers and Electro-Optics (CLEO)* (2013)
25. B.B. Guan, R.P. Scott, C. Qin, N.K. Fontaine, T.H. Su, C. Ferrari, M. Cappuzzo, F. Klemens, B. Keller, M. Earnshaw, S.J.B. Yoo, Free-space coherent optical communication with orbital angular, momentum multiplexing/demultiplexing using a hybrid 3D photonic integrated circuit. *Opt. Express* **22**(1), 145–156 (2014)
26. N. Bozinovic, P. Kristensen, S. Ramachandran, Long-range fiber-transmission of photons with orbital angular momentum, in *2011 Conference on Lasers and Electro-Optics (CLEO)* (2011)
27. Y. Yan, J. Wang, L. Zhang, J.Y. Yang, I.M. Fazal, N. Ahmed, B. Shamee, A.E. Willner, K. Birnbaum, S. Dolinar, Fiber coupler for generating orbital angular momentum modes. *Opt. Lett.* **36**(21), 4269–4271 (2011)
28. Y. Yan, L. Zhang, J. Wang, J.Y. Yang, I.M. Fazal, N. Ahmed, A.E. Willner, S.J. Dolinar, Fiber structure to convert a Gaussian beam to higher-order optical orbital angular momentum modes. *Opt. Lett.* **37**(16), 3294–3296 (2012)
29. Y. Yan, Y. Yue, H. Huang, J.Y. Yang, M.R. Chitgarha, N. Ahmed, M. Tur, S.J. Dolinar, A.E. Willner, Efficient generation and multiplexing of optical orbital angular momentum modes in a ring fiber by using multiple coherent inputs. *Opt. Lett.* **37**(17), 3645–3647 (2012)
30. J. Leuthold, R. Hess, J. Eckner, P.A. Besse, H. Melchior, Spatial mode filters realized with multimode interference couplers. *Opt. Lett.* **21**(11), 836–838 (1996)
31. J. Leuthold, J. Eckner, E. Gamper, P.A. Besse, H. Melchior, Multimode interference couplers for the conversion and combining of zero- and first-order modes. *J. Lightwave Technol.* **16**(7), 1228–1239 (1998)
32. J. Zhou, P. Gallion, Mode demultiplexers for rectangular multimode waveguides. *IEEE Photonics Technol. Lett.* **27**, 371–374 (2015)
33. S. Golowich, Asymptotic theory of strong spin–orbit coupling in optical fiber. *Opt. Lett.* **39**(1), 92–95 (2014)
34. S. Ramachandran, P. Gregg, P. Kristensen, S.E. Golowich, On the scalability of ring fiber designs for OAM multiplexing. *Opt. Express* **23**(3), 3721–3730 (2015)
35. J.H. Zhou, J.J. Wu, Q.S. Hu, Tunable arbitrary unitary transformer based on multiple sections of multicore fibers with phase control. *Opt. Express* **26**(3), 3020–3036 (2018)
36. R. Ryf, M.A. Mestre, S. Randel, X. Palou, A.H. Gnauck, R. Delbue, P. Pupaiaikis, A. Sureka, Y. Sun, X. Jiang, R. Lingle, Combined SDM and WDM transmission over 700-km few-mode fiber, in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)* (2013)
37. Z.X. He, X. Li, M. Luo, R. Hu, C. Li, Y. Qiu, S.N. Fu, Q. Yang, S.H. Yu, Independent component analysis based channel equalization for 6×6 MIMO-OFDM transmission over few-mode fiber. *Opt. Express* **24**(9), 9209–9217 (2016)
38. L.W. Luo, N. Ophir, C.P. Chen, L.H. Gabrielli, C.B. Poitras, K. Bergmen, M. Lipson, WDM-compatible mode-division multiplexing on a silicon chip. *Nat. Commun.* **5** (2014)
39. N. Hanzawa, K. Saitoh, T. Sakamoto, T. Matsui, K. Tsujikawa, M. Koshiba, F. Yamamoto, Two-mode PLC-based mode multi/demultiplexer for mode and wavelength division multiplexed transmission. *Opt. Express* **21**(22), 25752–25760 (2013)

40. J. Wang, S.T. Chen, D.X. Dai, Silicon hybrid demultiplexer with 64 channels for wavelength/mode-division multiplexed on-chip optical interconnects. *Opt. Lett.* **39**(24), 6993–6996 (2014)
41. K. Igarashi, D. Souma, K. Takeshima, T. Tsuritani, Selective mode multiplexer based on phase plates and Mach-Zehnder interferometer with image inversion function. *Opt. Express* **23**(1), 183–194 (2015)
42. H. Huang, G. Milione, M.P.J. Lavery, G.D. Xie, Y.X. Ren, Y.W. Cao, N. Ahmed, T.A. Nguyen, D.A. Nolan, M.J. Li, M. Tur, R.R. Alfano, A.E. Willner, Mode division multiplexing using an orbital angular momentum mode sorter and MIMO-DSP over a graded-index few-mode optical fibre. *Sci. Rep.* **5** (2015)
43. J. Liñares, X. Prieto-Blanco, V. Moreno, C. Montero-Orille, D. Mouriz, M.C. Nistal, D. Barral, Interferometric space-mode multiplexing based on binary phase plates and refractive phase shifters. *Opt. Express* **25**(10), 10925–10938 (2017)
44. G. Labroille, B. Denolle, P. Jian, P. Genevaux, N. Treps, J.F. Morizur, Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion. *Opt. Express* **22**(13), 15599–15607 (2014)
45. A. Crespi, R. Osellame, R. Ramponi, D.J. Brod, E.F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, F. Sciarrino, Integrated multimode interferometers with arbitrary designs for photonic boson sampling. *Nat. Photonics* **7**(7), 545–549 (2013)
46. G.T. Campbell, O. Pinel, M. Hosseini, T.C. Ralph, B.C. Buchler, P.K. Lam, Configurable unitary transformations and linear logic gates using quantum memories. *Phys. Rev. Lett.* **113**(6) (2014)
47. J.F. Morizur, L. Nicholls, P. Jian, S. Armstrong, N. Treps, B. Hage, M. Hsu, W. Bowen, J. Janousek, H.A. Bachor, Programmable unitary spatial mode manipulation. *J. Opt. Soc. Am. A-Optics Image Sci. Vision* **27**(11), 2524–2531 (2010)
48. R. Tang, T. Tanemura, Y. Nakano, Integrated reconfigurable unitary optical mode converter using MMI couplers. *IEEE Photonics Technol. Lett.* **29**(12), 971–974 (2017)
49. M. Reck, A. Zeilinger, H.J. Bernstein, P. Bertani, Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **73**(1), 58–61 (1994)
50. D.A.B. Miller, Self-configuring universal linear optical component. *Photonics Res.* **1**(1), 1–15 (2013)
51. D.A.B. Miller, Self-aligning universal beam coupler. *Opt. Express* **21**(5), 6360–6370 (2013)
52. J. Hudgings, L. Molter, M. Dutta, Design and modeling of passive optical switches and power dividers using non-planar coupled fiber arrays. *IEEE J. Quantum Electron.* **36**(12), 1438–1444 (2000)
53. N.K. Fontaine, R. Ryf, J. Bland-Hawthorn, S.G. Leon-Saval, Geometric requirements for photonic lanterns in space division multiplexing. *Opt. Express* **20**(24), 27123–27132 (2012)
54. T.A. Birks, I. Gris-Sánchez, S. Yerolatsitis, S.G. Leon-Saval, R.R. Thomson, The photonic lantern. *Adv. Opt. Photonics* **7**(2), 107–167 (2015)
55. H. Zhang, N. Healy, S. Dasgupta, J.R. Hayes, M.N. Petrovich, D.J. Richardson, A.C. Peacock, A tuneable multi-core to single mode fiber coupler. *IEEE Photonics Technol. Lett.* **29**, 591–594 (2017)
56. T. Hayashi, T. Taru, O. Shimakawa, T. Sasaki, E. Sasaoka, Design and fabrication of ultra-low crosstalk and low-loss multi-core fiber. *Opt. Express* **19**(17), 16576–16592 (2011)
57. Y. Sasaki, H. Uemura, K. Takenaga, S. Nishimoto, T. Uematsu, K. Omichi, R. Goto, S. Matsuo, K. Saitoh, Multicore fiber-based mode multiplexer/demultiplexer, in *Next-Generation Optical Communication: Components, Sub-Systems, and Systems IV*, vol. 9389 (2015)
58. J. Pietzsch, Scattering matrix analysis of 3×3 fiber couplers. *J. Lightwave Technol.* **7**(2), 303–307 (1989)
59. Z.S. Eznaveh, J.E. Antonio-Lopez, J.C.A. Zacarias, A. Schülzgen, C.M. Okonkwo, R.A. Correa, All-fiber few-mode multicore photonic lantern mode multiplexer. *Opt. Express* **25**(14), 16701–16707 (2017)
60. Z.I. Borevich, S.L. Krupetskii, Subgroups of the unitary group that contain the group of diagonal matrices. *J. Sov. Math.* **17**, 1951–1959 (1981)
61. J.H. Zhou, H.Q. Pu, Q.Q. Wang, Orbital angular momentum mode sorting based on a hybrid radial-angular hybrid lens. *Opt. Express* **30**(6), 9703–9713 (2022)

62. A. Mair, A. Vaziri, G. Weihs, A. Zeilinger, Entanglement of the orbital angular momentum states of photons. *Nature* **412**(6844), 313–316 (2001)
63. X.Y. Fang, H.R. Ren, M. Gu, Orbital angular momentum holography for high-security encryption. *Nat. Photonics* **14**(2), 102–108 (2020)
64. G. Ruffato, M. Massari, F. Romanato, Multiplication and division of the orbital angular momentum of light with diffractive transformation optics. *Light-Sci. Appl.* **8** (2019)
65. M. Padgett, J. Arlt, N. Simpson, L. Allen, An experiment to observe the intensity and phase structure of Laguerre-Gaussian laser modes. *Am. J. Phys.* **64**(1), 77–82 (1996)
66. J.M. Hickmann, E.J.S. Fonseca, W.C. Soares, S. Chávez-Cerda, Unveiling a truncated optical lattice associated with a triangular aperture using light's orbital angular momentum. *Phys. Rev. Lett.* **105**(5) (2010)
67. M.E. Anderson, H.H. Bigman, L.E.E.D. Araujo, J.L. Chaloupka, Measuring the topological charge of ultrabroadband, optical-vortex beams with a triangular aperture. *J. Opt. Soc. Am. B-optical Phys.* **29**, 1968–1976 (2012)
68. G.C.G. Berkhout, M.P.J. Lavery, J. Courtial, M.W. Beijersbergen, M.J. Padgett, Efficient sorting of orbital angular momentum states of light. *Phys. Rev. Lett.* **105**(15) (2010)
69. M.N. O'Sullivan, M. Mirhosseini, M. Malik, R.W. Boyd, Near-perfect sorting of orbital angular momentum and angular position states of light. *Opt. Express* **20**(22), 24444–24449 (2012)
70. R. Sahu, S. Chaudhary, K. Khare, M. Bhattacharya, H. Wanare, A.K. Jha, Angular lens. *Opt. Express* **26**(7), 8709–8718 (2018)
71. Y.H. Guo, S.C. Zhang, M.B. Pu, Q. He, J.J. Jin, M.F. Xu, Y.X. Zhang, P. Gao, X.G. Luo, Spin-decoupled metasurface for simultaneous detection of spin and orbital angular momenta via momentum transformation. *Light-Sci. Appl.* **10**(1) (2021)
72. G. Ruffato, OAM-inspired new optics: the angular metalens. *Light-Sci. Appl.* **10**(1) (2021)
73. V.V. Kotlyar, A.G. Nalimov, S.S. Stafeev, L. O'Faolain, Single metalens for generating polarization and phase singularities leading to a reverse flow of energy. *J. Opt.* **21**(5) (2019)
74. M. Vinas, C. Dorronsoro, V. Gonzalez, D. Cortes, A. Radhakrishnan, S. Marcos, Testing vision with angular and radial multifocal designs using adaptive optics. *Vision Res.* **132**, 85–96 (2017)
75. J.T. Xin, C.Q. Gao, Y.D. Liu, C. Li, K.J. Dai, Q.X. Na, Generation of Bessel beams from a diffractive ring lens. *Opt. Commun.* **310**, 25–30 (2014)
76. A. Volyar, M. Bretsko, Y. Akimova, Y. Gorov, Digital sorting perturbed Laguerre-Gaussian beams by radial numbers. *J. Opt. Soc. Am. A-Optics Image Sci. Vision* **37**(6), 959–968 (2020)
77. A. Volyar, M. Bretsko, Y. Akimova, Y. Egorov, Measurement of the vortex spectrum in a vortex-beam array without cuts and gluing of the wavefront. *Opt. Lett.* **43**(22), 5635–5638 (2018)
78. N.A.F. Zambale, G.J.H. Doblado, N. Hermosa, OAM beams from incomplete computer generated holograms projected onto a DMD. *J. Opt. Soc. Am. B-Opt. Phys.* **34**(9), 1905–1911 (2017)
79. A.V. Volyar, Y.A. Egorov, Super pulses of orbital angular momentum in fractional-order spiroid vortex beams. *Opt. Lett.* **43**(1), 74–77 (2018)
80. M.V. Berry, Optical vortices evolving from helicoidal integer and fractional phase steps. *J. Opt. A-Pure Appl. Opt.* **6**(2), 259–268 (2004)

Chapter 7

Temporal Modes and Spatiotemporal Modes



Abstract Temporal modes and spatiotemporal modes. The chapter extends the theory of spatial beam propagation to the temporal domain, and the concept of temporal waveguides and temporal Fractional Fourier transform are proposed. Afterwards, a temporal signal processing tool, i.e., a temporal diffractive deep neural network is proposed. The concept is extended beyond the spatial domain or the temporal domain, but the 4-dimensional space-time. A new class of spatial-temporal beam is discussed and a spatial-temporal diffractive deep neural network is proposed in the end.

Space–time duality has been an everlasting topic in physics, which has inspired ingenious contributions. The analogy of the space and time in classic optics has led to significant discoveries in optical research [1, 2]. For examples, the well-known Kramers–Kronig relation, which arises from the causal temporal response of the material [3], has its spatial counterpart [4]; the spatial domain reflection and refraction can happen in the time domain as well [5]. This chapter extends the theory of spatial beam propagation to the temporal domain, and the concept of temporal waveguides and temporal Fractional Fourier transform are proposed. Afterwards, a temporal signal processing tool, i.e., a temporal diffractive deep neural network is proposed. The concept is extended beyond the spatial domain or the temporal domain, i.e., the 4-dimensional space–time. A new class of spatiotemporal beam is discussed and a spatial–temporal diffractive deep neural network is proposed in the end.

7.1 Comprehensive Study on the Concept of Temporal Optical Waveguides

Part of the section reprinted/adapted with permission from [Phys. Rev. A 93, 063847 (2016)] © The American Physical Society [6].

In this section, we extend the concept of spatial optical waveguide to the temporal domain demonstrating the feasibility of confining an optical pulse within a specific

time frame through the imposition of temporal index barriers. This confinement enables the pulse to propagate at a speed of the index change in the waveguide, and it will be behind the original optical pulse that propagates without the temporal index variations. This approach offers a novel strategy for enhancing the bandwidth of slow light technology and adjusting light speed dynamically. The temporal waveguide possess modes, which are the temporal waveforms maintaining their shapes during the propagation. In a single mode temporal waveguide, the pulse preserves its shape as the sole mode like an optical soliton. In a multimode temporal waveguide, multimode interference (MMI) effect exists, which has the capability to replicate a single pulse into multiple instances, presenting a potential avenue for all-optical signal processing applications.

7.1.1 Introduction

Traditionally, reflection and refraction have been viewed as spatial events which cause change of the light propagation direction at the interface between two media possessing distinct refractive [7]. Nevertheless, temporal variations on the refractive index can likewise induce modifications in the shape and frequency of light pulses [8–10]. When such a temporal boundary of refractive index variation is present, the phenomena of reflection and refraction can appear in the time domain [5, 11]. Physically, this temporal refractive index boundary can be achieved through the employment of a travelling wave index modulator or a nonlinear waveguide with pronounced cross-phase modulation effects [5]. As is recently demonstrated by Plansinis et al. [5], in a dispersive media, total internal reflection (TIR) can appear temporally.

Optical beam confining structure like optical waveguide depends on the concept of TIR, which confines the optical wave within the edges. However, until now, there has been no analogous concept established in the time domain. The revelation of the potential for temporal TIR opens up the intriguing possibility that temporal waveguides could indeed exist if one may create two temporal boundaries of refractive index variation that facilitate TIR between them.

The confinement of light within a temporal waveguide within a brief duration offers a promising approach to achieve optical pulse width control. This is due to the tunability of the temporal boundaries, as demonstrated in [5]. Moreover, the temporal boundary's propagation speed (the rate of refractive index variation) can be set different from the optical beam's propagation speed, and it enables the slowing down of the optical pulse from its natural velocity.

Beyond modulating the speed of light, the analogy drawn between spatial and temporal waveguide characteristics stimulates novel applications for temporal optical waveguides. For instance, a single-mode temporal waveguide, like an optical soliton, enables a specific beam waveform, i.e., the fundamental temporal mode profile, to propagate unaltered within a dispersive medium. Another example is the extension of spatial multimode waveguide characteristics to the temporal domain. This opens doors to temporal all-optical signal processing techniques, mirroring the widespread

utilization of spatial multimode interference effects for optical coupling [12] and all-optical signal processing like orthogonal transforms [13–15]. The extension of the multimode waveguide to the temporal domain enables many temporal all optical signal processing techniques. For example, one can realize pulse self-imaging, a process where a single temporal pulse is replicated into multiple identical copies. When multiple pulses transmit in such a temporal multimode interference (MMI) waveguide, their input and output amplitudes become interrelated through a transfer matrix [15]. By modulating discrete phases on the pulses, it becomes feasible to execute temporal all-optical discrete Fourier transforms or discrete cosine/sine transforms for optical orthogonal frequency division multiplexing (OOFDM) systems [14, 15].

It is noteworthy that the concept under discussion has similarities to soliton-induced pulse trapping [16, 17], where the solitary wave induces the refractive index alterations, thereby confining a pulse within a temporal span. Nevertheless, the novel concept of temporal waveguides is more general. The trapped pulse does not interplay with the solitary trap as in [16, 17] and therefore, its propagation does not alter the property of the temporal waveguide. Moreover, this generalization paves the way for novel phenomena, including the confinement of sinusoidal-shaped pulses between temporal boundaries characterized by abrupt refractive index changes, as well as the exploration of temporal multimode interference effects. Lastly, the realization of temporal waveguides is not constrained solely to optically induced refractive index variations; alternative approaches, such as those based on radio-frequency (RF) signal-induced index modulations, offer additional ways for implementation.

7.1.2 Theory

We begin with how the temporal refractive index boundaries are implemented in a single mode spatial waveguide with uniform z directional parameters [5]. The waveguide can either be a travelling wave electro optical modulator, wherein the index varies with respect to the propagation of a microwave signal [5], or a waveguide exhibiting pronounced nonlinearity, enabling the adjustment of a probe signal's refractive index via intense pumping through cross-phase modulation [5]. The single mode spatial waveguide possesses chromatic dispersion. In such a dispersive single mode waveguide with a temporally varying refractive index, the propagation of the optical beam can be described using the equation governing the slowly evolving envelope of the optical pulse [5].

$$\frac{\partial \tilde{A}}{\partial z} + \Delta\beta_1 \frac{\partial \tilde{A}}{\partial t} + i\frac{\beta_2}{2} \frac{\partial^2 \tilde{A}}{\partial t^2} = i\beta_B H(t - T_B) \tilde{A} \quad (7.1)$$

where $\tilde{A}$ is the slowly varying envelope of the optical pulse, β_2 is the waveguide chromatic dispersion, $\Delta\beta_1 = \beta_1 - 1/V_B$, β_1 is the inverse of the group velocity and V_B is the temporal boundary propagation speed, $\beta_B = k_0 \Delta n$ is the propagation

constant variation caused by the index variation starting from $t = T_B$, T_B is the delay between the signal pulse and the index variation, and H is the Heviside function, which is 1 if $t > 0$ and 0 otherwise [5]. It is noteworthy that the dispersion effect on the controlling signal can affect the sharp boundary defined by the Heaviside function, transforming it into a boundary with a gradually varying index profile. Nevertheless, we can employ dispersion engineering techniques to design the waveguide in such a way that minimizes the dispersion impact on the controlling signal while enhancing it significantly for the signal of interest. By doing so, the Heaviside function remains a valid and useful approximation throughout our analysis.

It was shown in [5] that in case the following formula is valid, TIR will occur at the temporal boundary $t = T_B$,

$$\sqrt{2\beta_B\beta_2} > \Delta\beta_1 \quad (7.2)$$

If the two temporal boundaries for TIR are established, it is a logical conclusion that the optical pulse will undergo reflections between these boundaries, and thereby they are confined within the specified interval. To incorporate the second temporal boundary into our analysis, it is necessary to modify Eq. (7.1) accordingly.

$$\frac{\partial \tilde{A}}{\partial z} + \Delta\beta_1 \frac{\partial \tilde{A}}{\partial t} + i\frac{\beta_2}{2} \frac{\partial^2 \tilde{A}}{\partial t^2} = i\beta_B(H(t - T_B) + H(-t - T_B))\tilde{A} \quad (7.3)$$

The two boundaries are presumed to be situated at $-T_B$ and T_B , respectively, and their realization can be achieved by modulating the microwave signal or the intensity of the strong pump, which alters the refractive index. For the purpose of our analysis, the following change of variable is employed:

$$\tilde{A} = \exp\left(i\frac{\Delta\beta_1}{\beta_2}t\right)A \quad (7.4)$$

One may convert Eq. (7.3) to the following equation

$$\frac{\partial A}{\partial z} + i\frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} = i\left(\beta_B(H(t - T_B) + H(-t - T_B)) - \frac{\Delta\beta_1^2}{2\beta_2}\right)A \quad (7.5)$$

It is noteworthy that β_2 does not necessarily have to be positive; however, in case β_2 is negative, the Heaviside function on the right-hand side of the equation should be modified as $H(t + T_B) - H(t - T_B)$. It is crucial to acknowledge that Eqs. (7.3) and (7.5) produce identical results, yield equivalent outcomes, as evidenced by numerical simulations. Consequently, all simulations and analyses presented in this work will be based on Eq. (7.5).

Equation (7.5) is similar to the wave equation for the spatial optical waveguide adopting the paraxial approximation. To derive the analytical solution for the novel temporal waveguide modes, we presume a fast z directional oscillation, which suggests that the mode can be formulated as $A = \exp(j\beta z)B(t)$. Substituting it into

Eq. (7.5), we have:

$$\begin{aligned} \frac{\partial^2 B}{\partial t^2} + \frac{2\left(\beta + \frac{\Delta\beta_1^2}{2\beta_2}\right)}{\beta_2} B &= 0 \quad (-T_B < t < T_B) \\ \frac{\partial^2 B}{\partial t^2} + \frac{2\left(\beta + \frac{\Delta\beta_1^2}{2\beta_2} - \beta_B\right)}{\beta_2} B &= 0 \quad (t < -T_B \text{ or } t > T_B) \end{aligned} \quad (7.6)$$

Without loss of generality, β_2 is assumed positive. Given that the modes are restricted to the temporal interval $[-T_B, T_B]$, the function B must exhibit oscillatory behavior within this interval and decay rapidly outside it. Therefore, B can be expressed in the following forms:

$$\begin{aligned} B &= \begin{cases} A \sin(k_1 t) + B \cos(k_1 t) & -T_B < t < T_B \\ C \exp(k_2(t + T_B)) & t < -T_B \\ D \exp(-k_2(t - T_B)) & t > T_B \end{cases} \\ k_1 &= \sqrt{\frac{2\left(\beta + \frac{\Delta\beta_1^2}{2\beta_2}\right)}{\beta_2}} \\ k_2 &= \sqrt{\frac{2\left(\beta_B - \left(\beta + \frac{\Delta\beta_1^2}{2\beta_2}\right)\right)}{\beta_2}} \end{aligned} \quad (7.7)$$

Since the second order derivative $\frac{\partial^2 B}{\partial t^2}$ exists, the first order derivative $\frac{\partial B}{\partial t}$, and the function B should be continuous at the boundaries. Therefore, we may derive:

$$\begin{aligned} 2k_1 T_B &= 2 \tan^{-1}\left(\frac{k_2}{k_1}\right) + m\pi \\ k_1^2 + k_2^2 &= \frac{2\beta_B}{\beta_2} \end{aligned} \quad (7.8)$$

where m is an integer starting from 0. By solving the eigenvalue equation, i.e. Eq. (7.8), one can determine the corresponding field distribution (by evaluating k_1 and k_2) and the propagation constant β for each mode. The cut-off condition for the m th mode can then be derived:

$$2\sqrt{\frac{2\beta_B}{\beta_2}} T_B < m\pi \quad (7.9)$$

Similar to the spatial waveguide, the temporal waveguide can be designed to be single mode, which enables one specific shape of pulse to be contained within the interval, while other portion of the optical pulse dissipates. We may conclude the single mode condition for a temporal waveguide by setting $m = 1$ in Eq. (7.9).

7.1.3 Numerical Results and Discussions

It is noteworthy that the mode profile inside the temporal waveguide should be $\tilde{A}(t) = A(t) \exp\left(i \frac{\Delta\beta_1}{\beta_2} t\right)$. Given that $A(t)$, as is depicted in Eq. (7.7), varies relatively slowly when considered as the mode's profile, the inclusion of the high-frequency term $\exp\left(i \frac{\Delta\beta_1}{\beta_2} t\right)$ makes the mode solution of Eq. (7.3) a fast oscillation function with the parameters stated in [5] (i.e. $\beta_2 = 0.005 \text{ ps}^2/\text{m}$, $\Delta\beta_1 = 0.1 \text{ ps/m}$). This rapid oscillation could potentially make Eq. (7.3) or Eq. (7.5) invalid unless that the higher order dispersion terms are taken into account. In order to physically realize the concept of temporal waveguides, it is necessary to employ a waveguide or fiber characterized by an enhanced chromatic dispersion coefficient β_2 or a smaller $\Delta\beta_1$. This adjustment diminishes the oscillation frequency of the rapid oscillation term $\exp\left(i \frac{\Delta\beta_1}{\beta_2} t\right)$. Physically, this can be achieved through the utilization of either a dual concentric core waveguide design [18, 19], or a photonic crystal-based waveguide structure [20, 21] to enhance the chromatic dispersion. Notably, experimental reports in [19] indicate a normal dispersion value close to $2 \text{ ps}^2/\text{m}$, which we adopt in this study without loss of generality. Assuming that $\Delta\beta_1$ remains at 0.1 ps/m , the phase term $\exp\left(i \frac{\Delta\beta_1}{\beta_2} t\right)$ is interpreted as a carrier frequency shift of -50 GHz , which is a value that can keep Eqs. (7.3) and (7.5) valid without the necessity to incorporate higher-order dispersion terms. We investigate the feasibility of physically manifesting temporal waveguides by setting $\beta_2 = 2 \text{ ps}^2/\text{m}$, $\beta_B = 1.4 \text{ m}^{-1}$, $\Delta\beta_1 = 0.1 \text{ ps/m}$, which are from [5] with the exception of β_2 . We achieve single-mode and dual-mode operation by setting $2T_B = 2 \text{ ps}$ and $2T_B = 3 \text{ ps}$ (the single mode threshold for $2T_B$ is 2.6551 ps), respectively, as is demonstrated in Figs. 7.1 and 7.2. The initial pulse is a Gaussian pulse with a width of 6 ps , and its amplitude is normalized to unity for peak intensity. As the temporal waveguide operates in the linear regime and the signal does not interplay with the pump, the pulse amplitude's value does not influence the waveguide's characteristics. Thus, normalization does not alter the pulse's shape at the waveguide's output. In the single-mode waveguide, the pulse transforms into the fundamental mode, propagating as a solitary wave along the z -axis. In the dual mode temporal waveguide, the interplay between the first and second modes gives rise to two oscillating peaks along the z -axis.

Alterations in the temporal width of the waveguides give rise to the corresponding changes in the modal behavior. The details of the modal characteristics of the first four modes are demonstrated in Fig. 7.3, which tracks the evolution of the mode eigenvalue k_1 in relation to variations in the waveguide's width.

One may note that the single mode condition, i.e. Eq. (7.9), does not contain $\Delta\beta_1$, while the TIR condition includes $\Delta\beta_1$ (Eq. (7.2)). This suggests that the single-mode condition can be met even when the TIR condition is not satisfied.

One may compute the time delay induced by the temporal waveguide as

$$z\Delta\beta_1 = z(\beta_1 - 1/V_B) \quad (7.10)$$

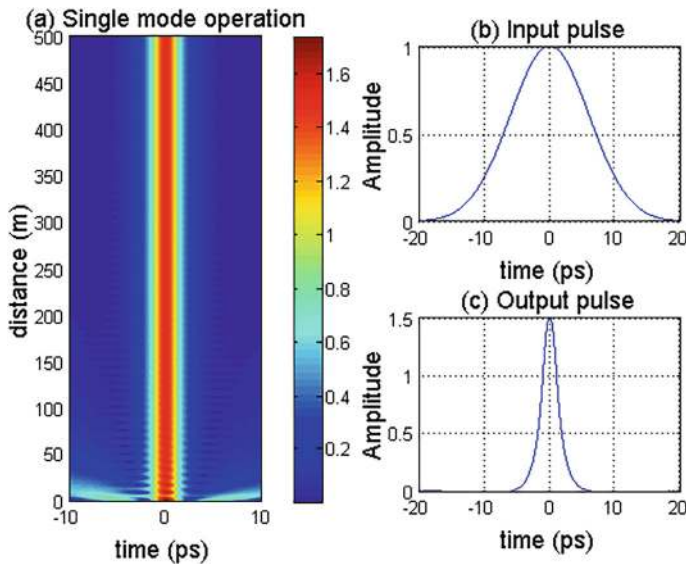


Fig. 7.1 Single mode temporal waveguide. **a** The amplitude evolution along the propagation distance. **b** The input pulse at $z = 0$ m. **c** The output pulse at $z = 500$ m. One may conclude from the Figure that the input pulse is changed by energy radiation from the two temporal waveguide boundaries. The stable sinusoidal temporal mode appears at the output. This Figure is reprinted/adapted with permission from [Phys. Rev. A 93, 063847 (2016)] © The American Physical Society [6]

which indicates that light within the waveguide propagates at a speed determined by the index boundaries, effectively slowing it down. This offers a novel perspective on light speed manipulation, allowing signals to travel relative to the speed of the index-controlling signals. Importantly, this approach aims to broaden the slow light bandwidth and enhance tunability, complementing existing methods rather than competing with them on the amount of speed reduction. On the one hand, photonic crystal [22] or cold atom vapors [23] for instance, often yield narrow slow light bandwidths. By leveraging the temporal waveguide concept and placing the pump at a wavelength corresponding to low velocities, we can significantly expand this bandwidth. On the other hand, since pulse speed is solely dependent on the speed of the controlling signal, high tunability can be achieved by adjusting the pump wavelength, as group velocity varies with wavelength [21].

Analogous to spatial waveguides, temporal waveguides also operate in a multi-mode regime. In spatial multimode waveguides, sufficient mode propagation triggers Multimode Interference (MMI), transforming an input image into multiple duplicates with different phases [13], and it is found that such a phenomenon also exists in the temporal waveguide. Remarkably, this MMI effect also occurs in the temporal domain, which is an unreported phenomenon with promising applications. It enables, for instance, the generation of multiple identical temporal signal copies, valuable for signal creation. In quantum optics, this concept could facilitate the production

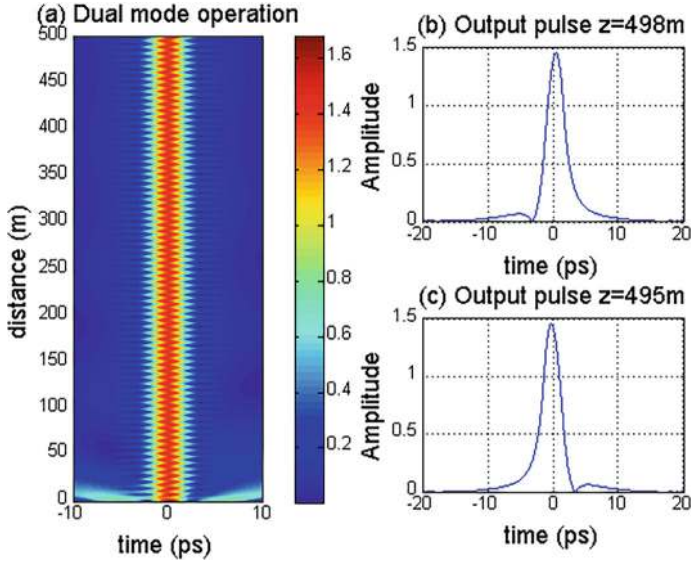


Fig. 7.2 Dual mode temporal waveguide. **a** The amplitude evolution along the propagation distance. **b** The output pulse at $z = 498$ m. **c** The output pulse at $z = 495$ m. The initial pulse shape of this figure is the same as the one shown in Fig. 7.1b. It can be seen from the figure that the initial pulse is reshaped during the propagation. The periodic pattern of energy distribution along the propagation path underscores the interplay between the two temporal waveguide modes. This Figure is reprinted/adapted with permission from [Phys. Rev. A 93, 063847 (2016)] © The American Physical Society [6]

of single photon sequence with an ultra short pulse. Furthermore, by modulating discrete phases on the pulses, temporal all-optical orthogonal transforms can be realized through the temporal MMI effect, facilitating advanced signal processing.

In scenarios where the time window width is substantial, the effective width of the temporal waveguide will be:

$$T_{eff} = 2T_B + 2 \sqrt{\frac{\beta_2}{\left(\beta_B - \left(\beta_1 + \frac{\Delta\beta_1^2}{2\beta_2}\right)\right)}} \quad (7.11)$$

where the last term stands for the Goos-Hanchen shift occurring at the two boundaries. Notably, during calculations, we adopt the propagation constant of the fundamental modum β_1 instead of those of higher-order modes, as the propagation constants tend to be a close value in multimode waveguides.

Within the effective time window, the optical pulse should obey the following equation:

$$\frac{\partial A}{\partial z} + i \frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} + \frac{i \Delta\beta_1^2}{2\beta_2} A = 0 \quad (7.12)$$

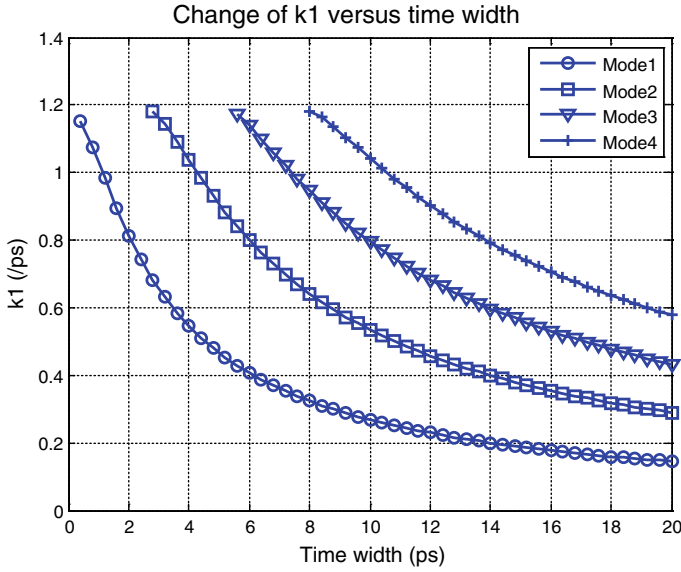


Fig. 7.3 Characterization of the mode eigen value k_1 with respect to the change of temporal waveguide width. The values of k_1 for the first four modes are shown. It can be seen that k_1 is almost inversely proportional to the temporal waveguide width. Each mode has a threshold except the fundamental mode. When the time width of the waveguide is below the threshold, the corresponding mode will not exist. This Figure is reprinted/adapted with permission from [Phys. Rev. A 93, 063847 (2016)] © The American Physical Society [6]

For the sake of simplifying the analytical representation of the approximate modes, we assume that the temporal waveguides start at $t = \sqrt{\frac{\beta_2}{\beta_B - \left(\beta_1 + \frac{\Delta\beta_1^2}{2\beta_2}\right)}}$, so that the mode distribution originates from $t = 0$ if the Goos-Hanchen shift is included. Importantly, adjusting the position of temporal boundaries does not alter any inherent properties of the temporal waveguide. Within such a framework, the modes can be approximated as:

$$A_m(z, t) = \sin\left(\frac{m\pi t}{T_{eff}}\right) \exp\left(i\left(\frac{\beta_2}{2}\left(\frac{m\pi}{T_{eff}}\right)^2 - \frac{\Delta\beta_1^2}{2\beta_2}\right)z\right) \quad (7.13)$$

To investigate the evolution of the optical pulse within the temporal waveguide, we decompose it into a sum of its propagation modes:

$$A(z, t) = \sum_{m=1}^{\infty} a_m A_m(z, t) = \sum_{m=1}^{\infty} a_m \sin\left(\frac{m\pi t}{T_{eff}}\right) \exp\left(i\left(\frac{\beta_2}{2}\left(\frac{m\pi}{T_{eff}}\right)^2 - \frac{\Delta\beta_1^2}{2\beta_2}\right)z\right) \quad (7.14)$$

where a_m represents the coefficient of each mode. Analogous to spatial MMI theory, temporal MMI occurs in the time domain at a specific propagation distance determined by:

$$z = L_N = \frac{2T_{\text{eff}}^2}{\pi\beta_2 N} \quad (7.15)$$

where N is the number of images.

To validate this, we simulate a multimode temporal waveguide with a time duration of 20 ps, bounded by $t = -10$ ps and $t = 10$ ps. The simulation employs parameters $\beta_2 = 2 \text{ ps}^2/\text{m}$, $\beta_B = 1.4 \text{ m}^{-1}$, $\Delta\beta_1 = 0.1 \text{ ps/m}$. By considering the Goos-Hanchen shift, the effective width should be 21.7 ps. Consequently, Eq. (7.15) predicts a propagation distance of 37.4 m for generating four MMI images, which is fine-tuned to 39 m after optimization. The input signal is a Gaussian pulse centered at $t = -8.13$ ps with the pulse width of 1 ps and it is also shifted in the carrier frequency by -50 GHz (i.e. multiplied by the phase term of $\exp(i\frac{\Delta\beta_1}{\beta_2}t)$). Four temporally separated images, each with a temporal spacing of $T_{\text{eff}}/4$, emerge at the output, confirming temporal MMI. The pulse's evolution is illustrated in Fig. 7.4.

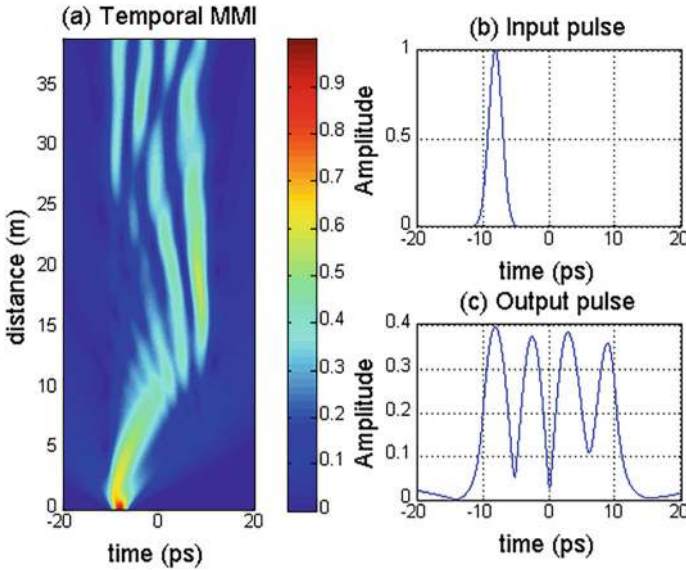


Fig. 7.4 Temporal multimode interference. **a** The amplitude evolution along the propagation distance. **b** The input pulse at $z = 0$ m. **c** The output pulse at $z = 39$ m. It can be seen from the figure that the interaction between the multiple modes have converted the input pulse into four duplicated pulses with identical shape. This Figure is reprinted/adapted with permission from [Phys. Rev. A 93, 063847 (2016)] © The American Physical Society [6]

7.1.4 Summary

In conclusion, we have introduced a novel concept of temporal waveguides, which mirrors its spatial counterpart in time domain. The boundaries of these temporal waveguides confine light beams, analogous to how spatial waveguides restrict light in space. Within waveguides or fibers exhibiting high dispersion, it becomes technically feasible to realize single-mode temporal waveguides, facilitating the propagation of pulses with stable waveforms like optical solitons. These solitary waves not only maintain their shape but also exhibit delays compared to pulses unconfined by temporal waveguides, thereby offering a means to modulate the speed of light. Additionally, the concept of temporal waveguides is equally valuable in the multimode regime, enabling the replication of input pulses into multiple identical copies. This proposed concept holds significant potential both theoretically and practically.

7.2 Fractional Fourier Transform Based on a Temporal Gradient-Index Lens

Part of the section reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24].

After taking analogy between space and time, it is worth introducing the tool for temporal signal processing. In this section, a novel approach to realize all optical fractional Fourier transform through a high-nonlinear fiber and an optical pump is discussed. Based on a gradient index (GRIN) time lens which is realized through cross-phase modulation by a temporal quadratic pump pulse, it is possible to achieve all optical fractional Fourier transform. The order of the fractional Fourier transform varies along the fiber propagation distance. In a fixed fiber, it can be adjusted by simply changing the peak power of the pump pulse.

7.2.1 Introduction

The Fractional Fourier Transform (FrFT), introduced by Namias in the 1980s for the exploration of quantum mechanics [25], has found versatile applications in signal processing [26–28]. This has sparked significant demand for its full optical implementation to enhance processing speed. It has been established that FrFT can be achieved through optical lenses [29–31], either the conventional spherical lenses or the gradient index (GRIN) lens. Numerous other optical FrFT implementation methods have been suggested. For instance, Cincotti [32] proposed the use of Array Waveguide Gratings (AWGs) for FrFT, while S. Weimann and colleagues recommended realizing discrete FrFT on a coupled waveguide array [33].

Temporal FrFT, analogous to the spatial FrFT mentioned above, holds particular importance in temporal signal processing. Like its spatial counterpart, temporal FrFT is accomplished through a combination of a quadratic phase generator (like a spatial lens) and dispersive elements (similar to free space propagation) [34, 35]. However, existing schemes are either difficult to adjust [34] or relatively complex in realization [35]. Besides these discrete time lenses, GRIN time lenses, which have been proposed for optical signal processing [36], have not been considered for FrFT implementation.

In this section, we introduce an innovative method for fully optical temporal FrFT implementation utilizing GRIN time lenses. These lenses can be easily created through the cross-phase modulation (XPM) effect in optical fibers [36]. As demonstrated in the subsequent analysis, the output wave, shaped by the GRIN time lenses formed by the quadratic pump pulse, transforms into the FrFT of the input wave. Notably, the order of FrFT evolves gradually with the propagation distance. Furthermore, as is explored in the upcoming section, the order of FrFT can be adjusted by varying the pump peak power, enhancing the flexibility of our proposed approach.

7.2.2 Theory

The schematic diagram of the GRIN time lens is presented in Fig. 7.5 [36]. In this setup, both the signal and the pump propagate in the same direction. A coupler, such as a WDM coupler, merges the temporal quadratic pump with the signal when the pump and signal operate at different wavelengths. Following this, a highly nonlinear fiber is introduced, within which the GRIN time lens occurs.

The propagation of the signal through the high nonlinearity optical fiber, influenced by the GRIN time lenses formed via the XPM effect, can be described by [36]:

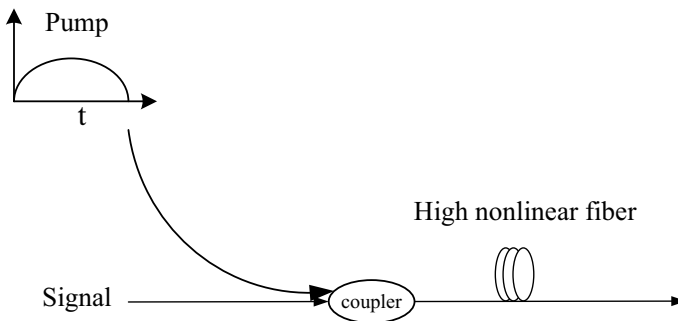


Fig. 7.5 The schematic of the GRIN time lens. This figure is reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24]

$$i \frac{\partial A}{\partial z} - \frac{\beta_2}{2} \frac{\partial^2 A}{\partial t^2} + 2\gamma P_p \left(1 - \frac{t^2}{\tau_p^2}\right) A = 0, \quad (7.16)$$

where A denotes the signal amplitude, β_2 represents the chromatic dispersion coefficient, γ is the nonlinear coefficient, P_p signifies the peak pump power, and τ_p indicates the half duration of the quadratic pump pulse. For simplicity, fiber loss is neglected, which is a reasonable assumption for shorter fibers [36]. Additionally, self-phase modulation (SPM) is disregarded by assuming a dominant pump and a weak probe signal.

The solution to Eq. (7.16) is expressed as [36]

$$\begin{aligned} A(z, t) &= \exp\left(i2\gamma P_p z - i\frac{\alpha z}{2}\right) \sum_{l=0}^{\infty} b_l \varphi_l(t) \exp(-il\alpha z), \\ \varphi_l(t) &= (\sigma_\tau 2^l l! \sqrt{\pi})^{-1/2} H_l\left(\frac{t}{\sigma_\tau}\right) \exp\left(-\frac{t^2}{2\sigma_\tau^2}\right), \\ b_l &= \int \varphi_l(u) f(u) du. \end{aligned} \quad (7.17)$$

where H represents the Hermite function, l denotes the order of the Hermite function, a_l is the coefficient, while α and σ_τ are the scaling parameters defined as [36]:

$$\begin{aligned} \alpha &= \sqrt{-\frac{\beta_2 \gamma P_p}{\tau_p^2}}, \\ \sigma_\tau &= \sqrt{-\frac{\beta_2}{\alpha}}. \end{aligned} \quad (7.18)$$

To facilitate further analysis, we apply the transformation:

$$u = \frac{t}{\sigma_\tau}. \quad (7.19)$$

FrFT has an inherent property described by [37]

$$\begin{aligned} F^p(\varphi_l(u))(u) &= e^{-jp\frac{\pi}{2}l} \varphi_l(u), \\ F^p(f(u'))(u) &= \sum_l^{\infty} b_l e^{-jp\frac{\pi}{2}l} \varphi_l(u). \end{aligned} \quad (7.20)$$

where F stands for the FrFT, and p represents the order of the FrFT. When $p = 1$, it corresponds to the standard Fourier Transform (FT). Upon comparing Eqs. (7.17 and 7.20), one has:

$$A(z, \sigma_\tau u) = \frac{1}{\sigma_\tau^2} \exp\left(i2\gamma P_p z - i\frac{\alpha z}{2}\right) F^{-\alpha z/(\frac{\pi}{2})}(f(A(0, u)))(u) \quad (7.21)$$

This suggests that $A(z, t)$ is essentially the FrFT of $A(0, t)$ with the scaling parameter σ_τ . Furthermore, the product of the parameter α and the propagation distance z correlates with the order of the FrFT, i.e., p :

$$p = \frac{2\alpha z}{\pi} = \frac{2z}{\pi} \sqrt{-\frac{\beta_2 \gamma P_p}{\tau_p^2}} \quad (7.22)$$

which indicates that the FrFT order evolves along the fiber's length. For a fiber with predetermined length, chromatic dispersion coefficient, and nonlinear coefficient, the order can be adjusted by modifying the pump pulse's peak power P_p and the pulse duration τ_p .

7.2.3 Results and Discussions

Numerical simulations have been carried out to validate the proposed approach utilizing GRIN time lenses for the full optical implementation of FrFT. The parameters and mathematical framework for the GRIN time lenses align with those discussed in [36]. To aid reader comprehension, we provide a brief overview: The chromatic dispersion coefficient $\beta_2 = -10 \text{ ps}^2/\text{km}$, the nonlinear coefficient $\gamma = 20 \text{ W}^{-1}/\text{km}$, the pump pulse peak power $P_p = 1 \text{ W}$, and the pulse duration $\tau_p = 16 \text{ ps}$. Although the simulated fiber specifications originate from [36], commercially available highly nonlinear fibers with anomalous dispersion can facilitate the realization of the GRIN time lens. In case that a standard single-mode fiber (SSMF) is utilized, the required pump power for FrFT execution would be approximately 17 W. The numerical simulations were performed using the finite difference (FD) method, incorporating perfect matching layers (PML) on both boundaries.

Without loss of generality, the input pulse is presumed to be a rectangular pulse, lasting 8 ps with a peak power of 1 mW, as is illustrated in Fig. 7.6. The amplitude has undergone normalization, ensuring that the input pulse peak exhibits unit power.

As is shown in Eq. (7.22) of the aforementioned section, the order of FrFT increases as the propagation distance increases. The distance required to realize the order 1 FrFT ($p = 1$), i.e., the standard FT can be computed as $Z_{\text{ft}} = 0.8886 \text{ km}$. The pulses at different propagation distances $z = 0.25Z_{\text{ft}}$, $0.5Z_{\text{ft}}$, $0.75Z_{\text{ft}}$, and Z_{ft} are shown in Fig. 7.7, which correspond to the FrFT orders of 0.25, 0.5, 0.75 and 1. For comparison, the FrFT curves of input signal are computed from the mathematical definition [37]:

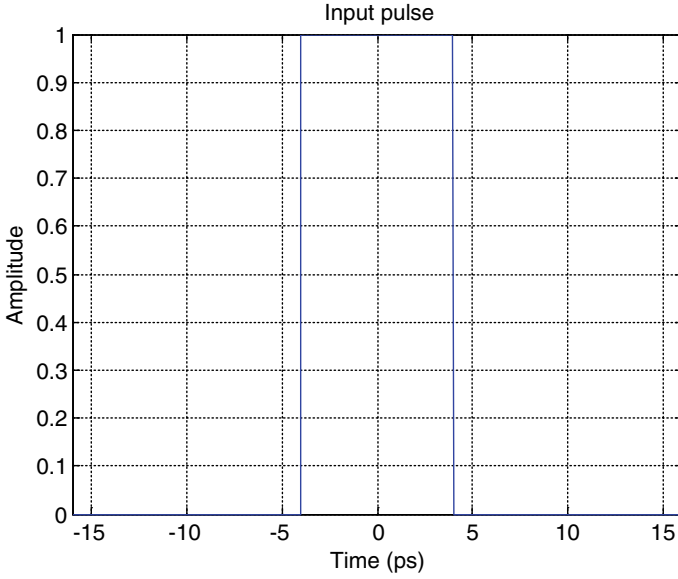


Fig. 7.6 The input rectangular pulse. The amplitude has been normalized such that the input pulse peak is with the unit power. This figure is reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24]

$$F^p(f(u'))(u) = \begin{cases} B_\theta \int_{-\infty}^{+\infty} \exp\left(j\frac{u'^2+u^2}{2} \cot \theta - \frac{juu'}{\sin \theta}\right) f(u') du' & \theta \neq k\pi \\ f(u) & \theta = 2k\pi \\ f(-u) & \theta = (2k+1)\pi \end{cases}, \quad (7.23)$$

where

$$\theta = \frac{p\pi}{2},$$

$$B_\theta = \sqrt{\frac{1 - j \cot \theta}{2\pi}}. \quad (7.24)$$

The close alignment between the curves indicates the realization of FrFT.

To illustrate the transformation more clearly, Fig. 7.8 depicts the variation of pulse amplitude with respect to propagation distance. Evidently, the self-image effect emerges at distances of $2Z_{\text{ft}}$ and $4Z_{\text{ft}}$, aligning precisely with the predictions made in [36]. The self-image effect can be understood as the reproduction of the original function by the FrFT when its order is 2 or 4, as is stated in the FrFT definition, Eq. (7.23).

As the order of FrFT evolves along the fiber propagation distance, it poses a challenge to capture different orders of FrFT from a single, fixed fiber. Optical tuning

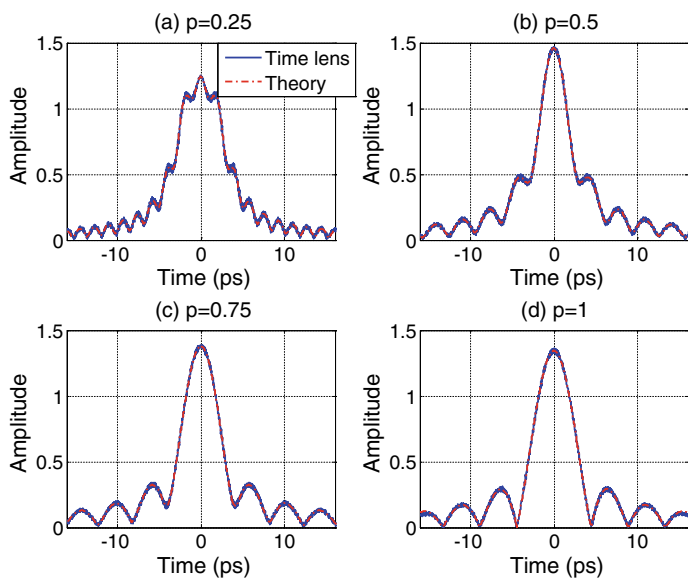


Fig. 7.7 The fractional Fourier transforms ($p = 0.25, 0.5, 0.75$ and 1) of the rectangular pulse in Fig. 7.6. $p = 1$ corresponds to the fiber propagation distance of 0.8886 km. The amplitude has been normalized with respect to the input pulse. This figure is reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24]

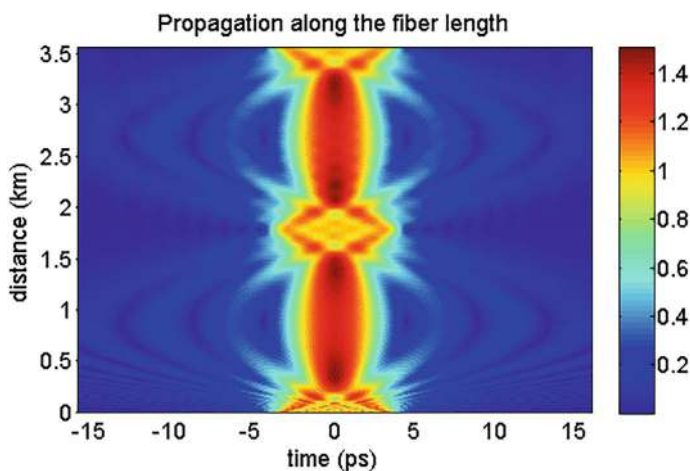


Fig. 7.8 The pulse amplitude evolution along the fiber length in the presence of the GRIN time lenses. The amplitude has been normalized with respect to the input pulse. This figure is reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24]

offers a more convenient approach. Equation (7.22) reveals that the order of FrFT can be modulated by adjusting the pump pulse peak power, enabling optical tunability. With the input pulse remaining constant, variations in pump peak power to 0.0625, 0.25, 0.5625, and 1 W were explored. According to Eq. (7.22), these adjustments correspond to FrFT orders of 0.25, 0.5, 0.75, and 1, respectively. The comparison between numerical simulations and theoretical FrFT calculations, are presented in Fig. 7.9. The remarkable alignment between the curves underscores that the proposed GRIN time lens can function as a tunable fractional Fourier transformer through simple adjustments to the peak power of the quadratic pump pulse. This innovative device requires only a high-nonlinear fiber and an optical pump to operate effectively.

Lastly, it is important to note that in the study of the temporal GRIN lens, it is assumed that the pump and signal are precisely synchronized, which may seem intuitive (see [36] and other references). However, synchronizing the two wavelengths can be challenging due to the group velocity mismatch between the pump and signal waves. This temporal mismatch can affect the performance of the optical fractional Fourier transformer, especially when processing short pulses.

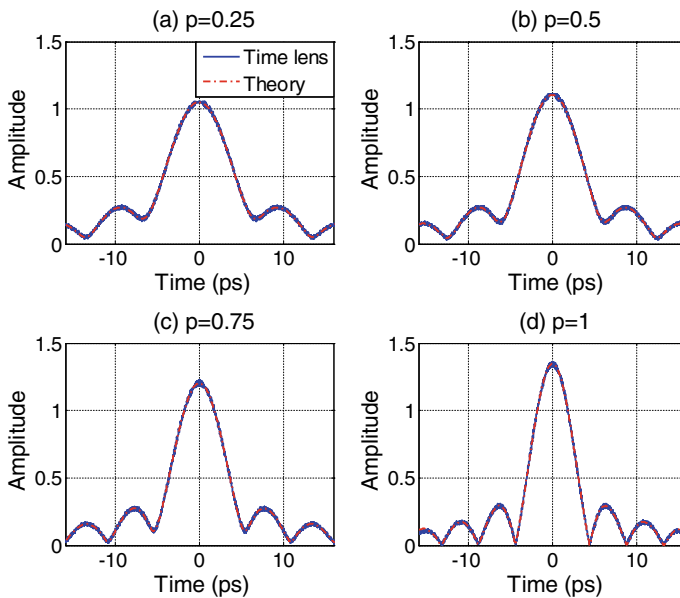


Fig. 7.9 The fractional Fourier transforms ($p = 0.25, 0.5, 0.75$ and 1) of the rectangular pulse in Fig. 7.1. The FrFT order is controlled by the different pump power with the values as 0.0625, 0.25, 0.5625 and 1 W respectively. The fiber propagation distance of 0.8886 km remains unchanged. The amplitude has been normalized with respect to the input pulse. This figure is reprinted/adapted with permission from [Opt Continuum 1, 1418–1424 (2022)] © The Optical Society of America [24]

7.2.4 Summary

We have introduced a novel approach to achieve an all-optical fractional Fourier transform utilizing an optical fiber. By creating a GRIN time lens through the XPM effect, wave propagation along the fiber can be interpreted as the fractional Fourier transform of the input pulse. The order of the FrFT varies with changes in fiber length. Additionally, the order of the FrFT can be adjusted by modifying the peak power of the pump pulse, indicating the feasibility of an all-optical tunable FrFT device. This concept has been validated through numerical simulations, which align closely with results obtained from FrFT theory. It is worth mentioning that the fractional transformer operates independently of pulse shape, despite the demonstration using an ideal pulse here.

7.3 Nonlinear Fourier Transform Receiver Based on a Time Domain Diffractive Deep Neural Network

Part of the section reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38].

Despite of the proposal of the temporal FRFT processor, a more general temporal processor is required. In this section, we introduce a diffractive deep neural network (D2NN) designed to differentiate temporal signals. As a demonstration of its capabilities, we apply the D2NN to classify inverse nonlinear Fourier transform (INFT) symbols. Unlike recently introduced D2NNs, our proposed network is based on optical fibers and operates in the time domain, rather than the spatial domain. The D2NN architecture consists of multiple cascaded dispersive elements and phase modulators. We propose an all-optical back-propagation algorithm to optimize the phase settings within the network. This fiber-based, time domain D2NN serves as a robust tool for signal conversion and recognition, particularly at the receiver side to optically recognize the INFT symbols. Following the symbol conversion by the D2NN, simple phase and amplitude measurements suffice to determine the correct symbol, eliminating the need for the time-consuming nonlinear Fourier transform (NFT) process. The versatility of this device extends its potential applications not only within the NFT transmission systems but also in various other areas that require all-optical time domain signal transformation and recognition, such as sensing, signal encoding and decoding, beam distortion compensation, and image recognition.

7.3.1 Introduction

NFT has emerged as a promising approach to overcome the limitations posed by fiber nonlinearity, thereby enhancing the transmission capacity of the system. The fundamental model describing wave propagation in optical fibers, namely, the nonlinear Schrödinger equation (NLSE), constitutes an integrable system. Beams converted by inverse NFT (INFT) serve as ideal information carriers, undergoing only phase changes in the nonlinear spectral domain during transmission [39–49]. Although theoretical analyses indicate a significant potential for increasing system capacity [42], designing an NFT receiver poses challenges due to the complexity involved in numerically realizing NFT.

Machine learning technology has seen rapid advancements and found applications in various domains such as computing, image processing, video and audio signal recognition, signal encoding and decoding, etc. [50]. This technology has also proven effective in achieving NFT decoding through symbol classification [43–49]. However, most machine learning implementations operate in the electronic domain, which can be time-consuming. All-optical machine learning aims to process data at the speed of light, reducing computational complexity. The diffractive deep neural network (D2NN) is considered as one of the recent breakthroughs in all-optical machine learning [51–58]. It employs cascaded free-space propagation and amplitude/phase modulation to create a tunable spatial beam transformer, enabling a wide range of all-optical spatial signal processing tasks [51–58]. Nevertheless, as the D2NN operates in the spatial domain, it isn't directly applicable for NFT decoding.

It is widely recognized that time and space share analogous properties, allowing concepts from the spatial domain to be translated into the temporal domain by substituting spatial elements with their temporal counterparts [1]. Therefore, we introduce an all-optical time-domain D2NN designed to classify INFT symbols. This is achieved by converting spatial propagation into temporal propagation through dispersive elements and adapting spatial amplitude or phase modulation to the time domain.

It is well-known that sequential dispersive propagation and phase modulation can be implemented to achieve time-domain unitary transforms. Previous studies have suggested that an all-optical time lens can be constructed using two dispersive elements and a quadratic phase modulator [59]. Similarly, an all-optical fractional Fourier transform can be implemented in this manner [35]. Additionally, an all-optical unitary transformer for quantum signal processing, realized through a series of dispersive elements and phase modulators, has been proposed and discussed [60]. While these groundbreaking studies are inspiring, they primarily focus on all-optical unitary transforms and have not explored signal recognition applications. Furthermore, existing research often relies on global search methods like the simulated annealing algorithm for phase tuning [60, 61], while more efficient gradient search techniques have not been explored due to the challenge of deriving closed-form expressions for gradients. Thus, there is a strong need for an effective phase tuning algorithm.

In this study, we present a novel time-domain D2NN, incorporating sequential dispersive elements and phase modulators, aimed at the automatic recognition of INFT symbols. Following the conversion of INFT symbols by the D2NN, simple phase and amplitude measurements aim to determine the correct symbol, eliminating the need for time-consuming NFT computations. Given that the time-domain D2NN is fiber-based, it can be seamlessly integrated into the NFT transmission system. During the training phase, we introduce an all-optical back-propagation (BP) algorithm which involves the inverse propagation of the conjugated error signal through the D2NN. By optically measuring this conjugated error signal, one can compute the gradient to adjust the phases layer by layer.

7.3.2 The Time Domain D2NN and All Optical Error Back-Propagation

The proposed time-domain D2NN consists of multiple layers, with each layer made up of a neuron layer implemented through an optical phase modulator (OPM), followed by a dispersive element (refer to Fig. 7.10a). The dispersive element can take the form of a chirped fiber Bragg grating (CFBG) (as shown in Fig. 7.10b), a segment of optical fiber (Fig. 7.10c), or other dispersive components, as illustrated in Fig. 7.10.

To describe the propagation within the time-domain D2NN, we utilize cascaded operators, represented by the equation:

$$\mathbf{U}_k = \mathbf{N}_k \mathbf{D}_k \mathbf{U}_{k-1}, \quad (7.25)$$

where $\mathbf{U}_k$ denotes the time-domain waveform after it propagates through the k th layer, $\mathbf{D}_k$ represents the propagation operator for the dispersive element, $\mathbf{N}_k$ stands for an operator describing the impact of the phase modulation. If the system is discretized in time, $\mathbf{U}_k$ becomes a vector, while $\mathbf{D}_k$ and $\mathbf{N}_k$ are represented as matrices. Particularly, $\mathbf{N}_k$ is a diagonal matrix where each diagonal element corresponds to a controllable neuron n_k^l . Those neurons form a neuron vector $\mathbf{n}_k^T$, expressed as $\exp(j\phi_k)$, and the relationship between the matrix and vector is given by $\mathbf{N}_k = \text{diag}(\mathbf{n}_k)$. This equation is similar to the one for the spatial-domain D2NN [53]. The final output is related to the input by:

$$\mathbf{U}_{N+1} = \mathbf{D}_{N+1} \left(\prod_{k=N}^1 \mathbf{N}_k \mathbf{D}_k \right) \mathbf{U}_0. \quad (7.26)$$

where $\mathbf{U}_0$ is the input signal and $\mathbf{U}_{N+1}$ is the output signal converted by the time domain D2NN.

It may facilitate understanding by comparing the D2NN concept to the split-step Fourier method (SSFM). Indeed, there are notable similarities between these two

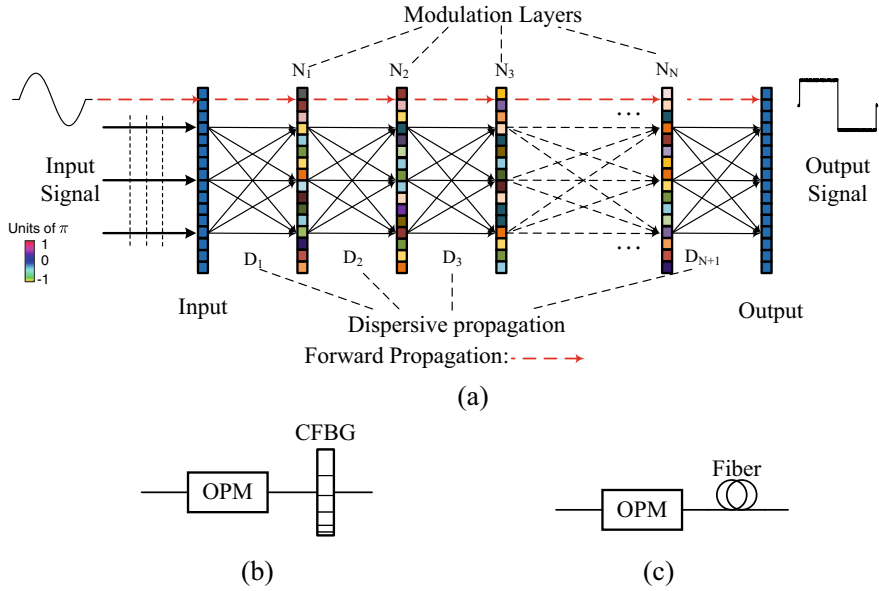


Fig. 7.10 Basic structure of the D2NN. **a** Basic schematic for the time domain D2NN, **b** the layer formed by a chirped fiber grating (CFBG) and the optical phase modulator (OPM) or **c** a piece of optical fiber and the optical phase modulator. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

approaches, both of which employ cascaded dispersive operators and phase operators. However, a key difference lies in the nature of the phase operators. In SSFM, the phase operators are nonlinear, meaning the phase depends on the amplitude of the input waveform during the nonlinear step. Meanwhile, in the D2NN, the phase operators are linear, with the phase remaining fixed (albeit learned through the subsequent BP algorithm) during wave propagation. Thus, while there are significant similarities between the two methods, there are also distinct differences.

The loss function is defined as

$$L = |\mathbf{U}_{N+1} - \mathbf{T}|^2, \quad (7.27)$$

where $\mathbf{T}$ represents the target waveform, incorporating both amplitude and phase information. However, if only the intensity is relevant, as in the case of the spatial domain D2NN [57], the loss function can be redefined accordingly. To adjust the neurons, it's necessary to evaluate the gradient of the loss function:

$$\frac{\partial L}{\partial \mathbf{n}_k^T} = (\mathbf{U}_{N+1} - \mathbf{T})^H \frac{\partial \mathbf{U}_{N+1}}{\partial \mathbf{n}_k^T},$$

$$\begin{aligned}
\frac{\partial \mathbf{U}_{N+1}}{\partial \mathbf{n}_k^T} &= \mathbf{D}_{N+1} \left(\prod_{p=N}^{k+1} \mathbf{N}_p \mathbf{D}_p \right) \frac{\partial \mathbf{N}_k}{\partial \mathbf{n}_k^T} \mathbf{D}_k \left(\prod_{q=k-1}^1 \mathbf{N}_q \mathbf{D}_q \right) \mathbf{U}_0 \\
&= \mathbf{D}_{N+1} \left(\prod_{p=N}^{k+1} \mathbf{N}_p \mathbf{D}_p \right) \text{diag} \left(\mathbf{D}_k \left(\prod_{q=k-1}^1 \mathbf{N}_q \mathbf{D}_q \right) \mathbf{U}_0 \right). \tag{7.28}
\end{aligned}$$

Thanks to the symmetrical property of the matrices:

$$\begin{aligned}
\mathbf{D}_k^T &= \mathbf{D}_k, \\
\mathbf{N}_k^T &= \mathbf{N}_k,
\end{aligned} \tag{7.29}$$

we have:

$$\begin{aligned}
\frac{\partial L}{\partial \mathbf{n}_k^T} &= \left(\mathbf{q}_k^f \odot \mathbf{p}_k^b \right)^T, \\
\mathbf{q}_k^f &= \left(\left(\mathbf{D}_k \prod_{q=k-1}^1 \mathbf{N}_q \mathbf{D}_q \right) \mathbf{U}_0 \right), \\
\mathbf{p}_k^b &= \left(\left(\prod_{p=k+1}^N \mathbf{D}_p \mathbf{N}_p \right) \mathbf{D}_{N+1} (\mathbf{U}_{N+1}^* - \mathbf{T}^*) \right),
\end{aligned} \tag{7.30}$$

where $\odot$ denotes the element-wise product. For the phase tunable neurons, we have $\mathbf{n}_k = \exp(j\varphi_k)$, with φ_k representing the phase vector. Therefore, one has:

$$\begin{aligned}
\frac{\partial L}{\partial \varphi_k^T} &= j \frac{\partial L}{\partial \mathbf{n}_k^T} \mathbf{N}_k + c.c. \\
&= 2\text{Re} \left(j \frac{\partial L}{\partial \mathbf{n}_k^T} \mathbf{N}_k \right) \\
&= 2\text{Re} \left(j \mathbf{p}_k^f \odot \mathbf{p}_k^b \right)^T, \\
\mathbf{p}_k^f &= \mathbf{N}_k \mathbf{q}_k^f = \left(\prod_{q=k}^1 \mathbf{N}_q \mathbf{D}_q \right) \mathbf{U}_0,
\end{aligned} \tag{7.31}$$

where c.c. stands for the complex conjugate. Equations (7.30 and 7.31) reveal that the gradient of the loss function can be computed in a backward manner. Notably, the vector $\mathbf{p}_k^f$ corresponds to the forward-propagating optical field after the k th layer, which can be directly obtained through forward optical propagation. Meanwhile, the vector $\mathbf{p}_k^b$ can be derived through all-optical back-propagation of the conjugated error field at the same position. To obtain the conjugated error wave, two steps are involved. Firstly, all-optical interference between the output signal of the DNN

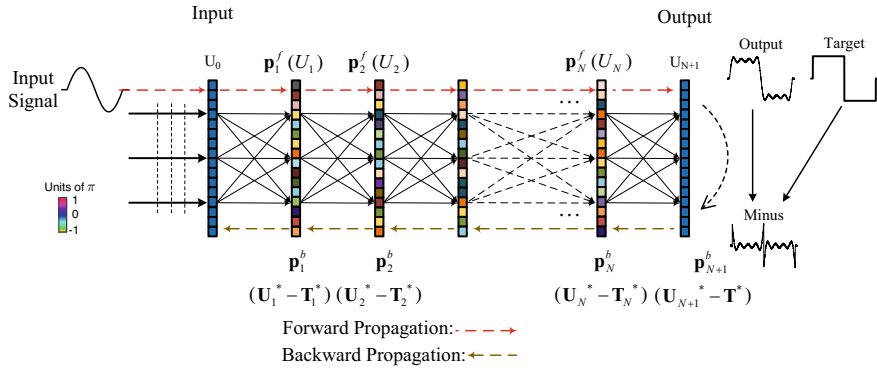


Fig. 7.11 Forward and backward propagation of the vectors $\mathbf{p}_k^f$ and $\mathbf{p}_k^b$. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

and the target waveform can be conducted, enabling the optical subtraction of the two waveforms. Secondly, the conjugated error signal can be generated through parametric amplification, which produces a high-fidelity phase-conjugated signal on the idler wave. The preceding analysis demonstrates the feasibility of implementing an all-optical BP algorithm, and a detailed illustration is provided in Fig. 7.11.

We may update the phase in each layer by implementing the gradient:

$$\varphi_k^{new} = \varphi_k^{old} - \mu \frac{\partial L}{\partial \varphi_k^{old}}, \quad (7.32)$$

where μ stands for the learning rate. After several iterations, the phases are adjusted to the optimal value.

7.3.3 Results and Discussions

In our simulation, the dispersive element exhibits a chromatic dispersion of 160 ps²/nm, equivalent to that of a 10 km-long fiber, as referenced in [60]. For simplicity, we assume a chirped fiber grating (CFBG) as the dispersive element, as mentioned in [60]. Within a 1 ns time frame, there are 128 neurons, indicating that the phase modulator can adjust the phase within approximately 8 ps intervals. It's important to note that each layer of the time domain DNN may contain a varying or consistent number of neurons; in our simulation, we assume uniform neuron counts across layers. Such phase modulation is feasible given the availability of 64 Gbaud coherent transceivers, which are capable of generating amplitude and phase modulated signals at 128 G samples per second. A single phase modulator and dispersive element constitute a D2NN layer.

NFT transmission is implemented on a polarization-multiplexed (Pol-Mux) system with a discrete NFT spectrum. This spectrum has two eigenvalues of $0.3j$ and $0.6j$, as is stated in [40]. These eigenvalues are modulated using QPSK signals, with radii of 0.14 and 5, respectively. A phase difference of $\pi/4$ is introduced during modulation for both eigenvalues, following the method outlined in [40]. The transmission system operates at a baud rate of 1 G Baud/s, implying that each symbol has a duration of 1 ns [40]. Although we have two polarizations and two eigenvalues for QPSK modulation, resulting in a potential total of 256 distinct symbols, many of these are similar in amplitude and phase due to NFT properties. Therefore, we identify 16 “independent” symbols for the Pol-Mux system, which differ in amplitude or phase distributions. These independent symbols are illustrated in Fig. 7.12. It’s worth noting that Fig. 7.12 only depicts power distributions, and some pulses may appear similar. However, they differ in phase distributions, as is demonstrated in Fig. 7.13 (e.g., compare symbol 1 and symbol 9 in Figs. 7.12 and 7.13). With various input symbols, the desired outputs are rectangular pulses in specific time slots. By measuring the power within these time slots, we can determine the symbol without time-consuming NFT.

T D2NN phase optimization is achieved by adjusting the phase based on the phase gradient computed from the forward propagation error vector $\mathbf{p}_k^f$ and the backward propagation error vector $\mathbf{p}_k^b$. Both error vectors are derived through all-optical propagation. During the training process, we use 16 classes of symbols with targets set as ones in corresponding time slots for both input and output. The proposed BP algorithm is applied over 10,000 iterations to obtain optimized phase distributions. The learning rate for this process is set at 0.001.

We consider two distinct transmission systems. The first features a shorter span of 40 km, whereas the second has a longer reach of 80 km, like the systems described in [40], with fiber lengths of 41.5 and 83 km, respectively. We assume to use the standard single-mode fibers for these spans, characterized by a chromatic dispersion of $16 \text{ ps}^2/\text{nm/km}$ and a nonlinear coefficient of $1.2 \text{ W}^{-1}/\text{km}$. To offset fiber attenuation, optical amplifiers are positioned after each span, with adjustable noise figures (NFs) to fine-tune the signal-to-noise ratio (SNR) at the output. The shorter fiber span facilitates a transmission system that closely aligns with an integrable system, whereas the longer span may exhibit propagation characteristics that diverge from those of a perfect integrable system. The propagation of INFT symbols within these optical links is simulated using the coupled Manakov equation.

The short span transmission operates at a relatively lower power level, with signals propagating across 5 spans covering a total distance of 200 km. Five optical amplifiers, each with a NF of 5 dB, are inserted after each span. Figure 7.14 illustrates the demultiplexing results for 16 independent signals. The D2NN consists of 14 layers, with each layer comprising 128 neurons per nanosecond. Given that the targets for the time domain D2NN conversion are rectangular pulses, the converted signals shown in Fig. 7.14 exhibit a stair-shape pulse pattern. This approach transforms the signal into an intensity-modulated format, simplifying recognition. As is previously discussed, one can determine the correct symbol by analyzing the power levels within specific time slots. For instance, the first sub-figure in Fig. 7.14 shows the converted pulse for

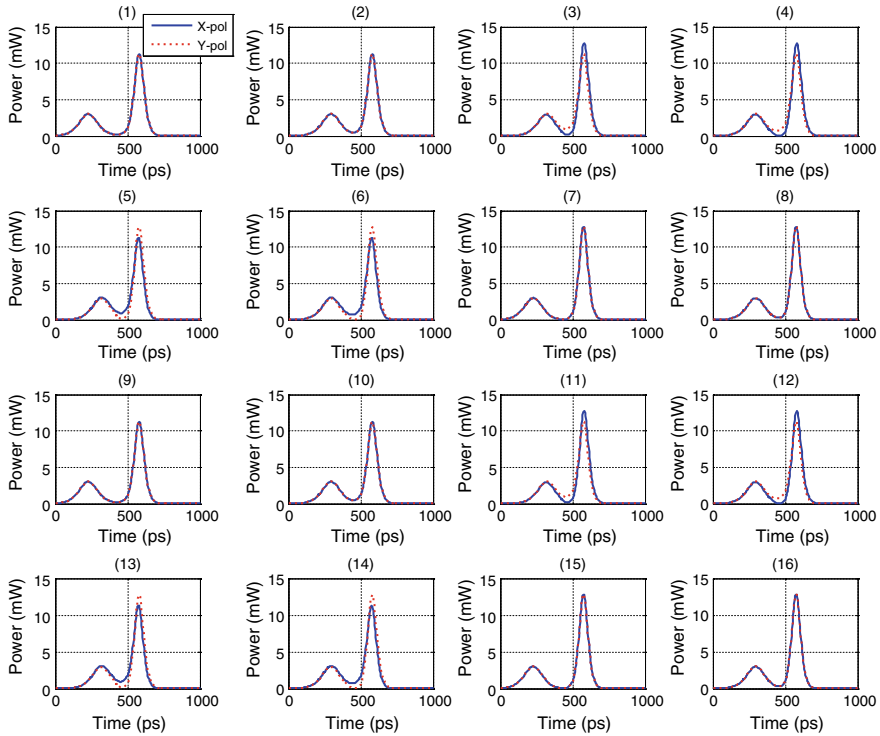


Fig. 7.12 The 16 independent pulses for the Pol-Muxed NFT systems with two discrete eigen values. The QPSK signals are modulated on the discrete nonlinear spectrum. This figure is reprinted/ adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

the first independent symbol from Fig. 7.12. This pulse displays the highest power concentration in the first time slot for both x and y polarizations. By comparing power levels across different time slots, one can deduce that this corresponds to the first symbol in Fig. 7.12 after undergoing conversion via the time domain D2NN.

While the 16 independent symbols can be effortlessly identified through power measurements, signals with identical amplitude and phase distributions for x and y polarizations (aside from a constant phase term) can be differentiated through relative phase measurements. Phases are obtained through coherent measurements of the signals. Since these signals are modulated using QPSK, there are four distinct phases for each polarization, resulting in 16 possible phase combinations for x–y polarizations within a single symbol class. Based on these phases, the received INFT symbols can be categorized into 256 unique types. Carrier frequency offset, phase offset estimation, and compensation are performed [46] before decision-making. Thus, INFT symbols can be optically recognized without relying on complex NFT processes. It's important to note that special attention should be given to potential

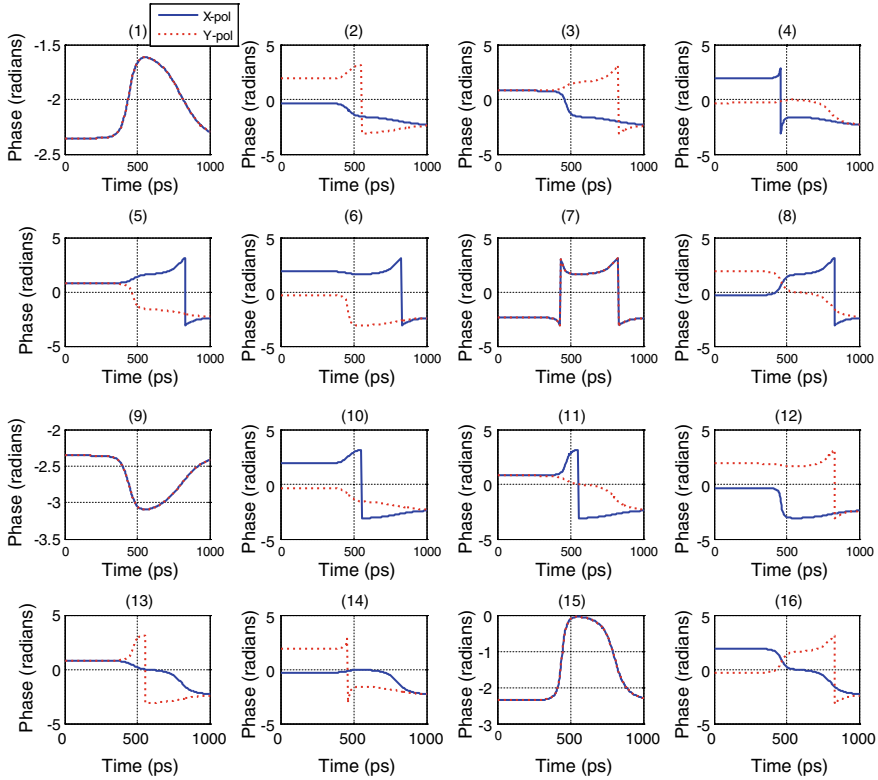


Fig. 7.13 The phase distributions of the 16 independent pulses in Fig. 7.3. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

phase noise from the laser source, and appropriate estimation and compensation algorithms should be employed [46].

A total of 256,000 symbols were transmitted in the 5-span system covering an overall length of 200 km for the evaluation of symbol error rates (SERs). Figure 7.15 illustrates the SER of the demultiplexed signals presented in Fig. 7.14, considering various signal-to-noise ratios (SNRs). Missing points in the graph indicate a zero SER for the corresponding 256,000 symbols. It's important to note that accurate SER estimation requires the counting of at least 100 errors; therefore, slight fluctuations in the figures might be attributed to numerical artifacts. The SNR can be adjusted by modifying the NFs of the optical amplifiers. Evidently, from the graph, the SER increases as the SNR decreases. Additionally, the SER is influenced by the number of neuron layers in the D2NN. A minimum of 7 layers is required for successful demultiplexing. When the layer count reaches 11, performance stabilizes, making it suitable for practical receivers.

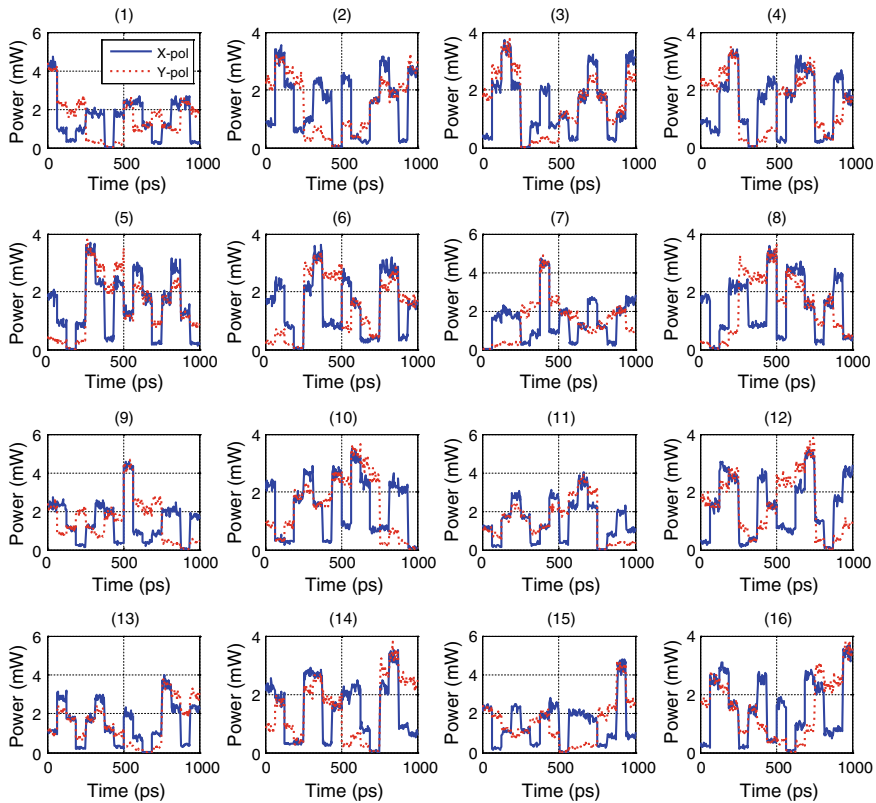


Fig. 7.14 The 16 independent pulses demultiplexing results with the fiber-based time domain D2NN. The targets are the rectangular pulses during the D2NN training. One may collect the powers within the time slots to determine which INFT symbol it should be. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

For comparison, we present NFT results obtained using the layer peeling method combined with the Newton method to derive discrete spectrum values. The NFT was performed with 128 samples per symbol, resulting in a significantly lower SER compared to the time domain D2NN results. (If 64 samples per symbol were used in NFT, as in [40], the SER performance would be notably poorer than that of the 11-layer D2NN.) However, considering the substantial computational savings during the decoding process, the proposed all-optical approach emerges as a promising alternative to NFT. Furthermore, NFT results deteriorate as imperfections accumulate during propagation, whereas the D2NN remains unaffected by such penalties. This advantage of D2NN will become evident in the propagation analysis of long-span systems, as is discussed below.

As is illustrated in Fig. 7.15, the 11-layer D2NN appears to offer an excellent balance between performance and system complexity, making it a suitable choice

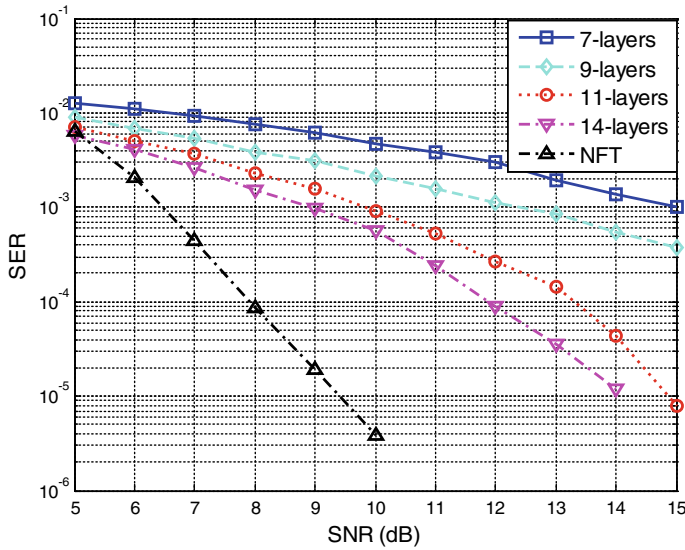


Fig. 7.15 The SER of the de-multiplexed signals with respect to different SNRs. The fiber span length is 40 km. The number of spans is 5. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

for analysis in the second transmission system with a fiber span of 80 km. To assess the influence of span number on SER, we maintain a constant SNR of 10 and 15 dB. The curves in Fig. 7.16 reveal that the SER generated by the time domain D2NN does not consistently rise with an increase in span number when the SNR remains unaffected. This behavior is attributed to the fact that D2NN's demultiplexing efficiency relies on the orthogonality between symbols. Meanwhile, nonlinear transformations like NFT disrupt symbol orthogonality, leading to performance degradation. When the overlap integral between symbols is minimal, D2NN is more likely to deliver effective demultiplexing outcomes. This overlap integral can fluctuate during NFT system propagation, explaining the irregular patterns observed in the SER curves relative to span number.

For comparative purposes, Fig. 7.16 also displays NFT results. Notably, the SER in NFT increases proportionally with the number of spans, aligning with experimental measurements reported in [40] for an NFT system. This trend occurs because imperfect fiber channels introduce attenuation, and as the number of spans rises, the fiber transmission system diverges from an integrable system. Consequently, waveforms deviate further from ideal INFT symbols after extended propagation distances. In such scenarios, the proposed D2NN-based receiver demonstrates superior performance.

Lastly, we investigate the impact of phase neuron precision. Assuming an SNR of 15 dB, Fig. 7.17 illustrates the scenarios where the phase modulators have phase resolutions of 1° and 5° , respectively. The SER curves reveal that there is almost

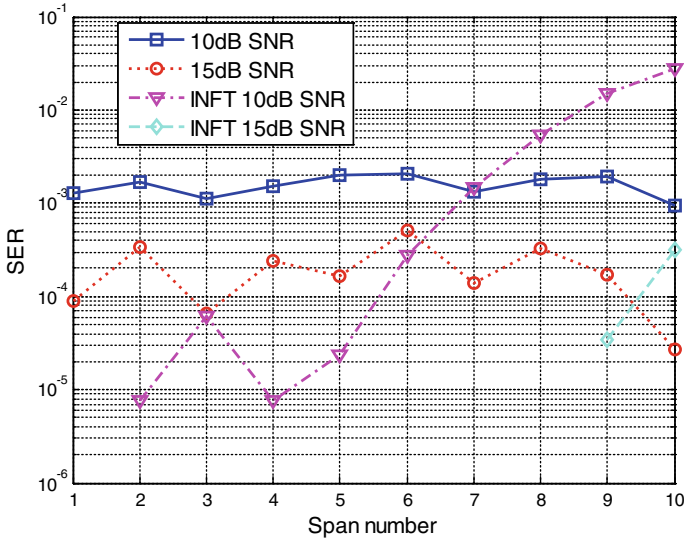


Fig. 7.16 The SER of the de-multiplexed signals with respect to fiber spans under the SNRs of 10 and 15 dB. The fiber span length is 80 km. A 11-layer D2NN is used for INFT symbol recognition. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

no difference between the results obtained by the D2NN with a 1-degree phase resolution and the D2NN without any phase error. Even the SER curve generated by the D2NN with a 5-degree phase resolution closely aligns with the other two curves. Given that the testing symbols differ from the training symbols (due to varying noise realizations), it's possible that the training optimum might not align with the testing optimum. Figure 7.17 shows slight performance enhancements achieved by the 5-degree precision D2NN at certain span numbers. Fortunately, these performance fluctuations remain minimal. The findings in Fig. 7.17 indicate that the proposed D2NN exhibits remarkable robustness against modulator phase errors.

It's crucial to note that performance is also influenced by the number of distinct symbol classes to be identified. In such cases, using a single eigenvalue yields better results since there are fewer independent symbol classes. However, when three eigenvalues are employed, more independent symbol classes emerge, leading to reduced performance.

Therefore, the introduced time domain D2NN has proven its effectiveness in INFT symbol recognition. The complex NFT is no longer necessary at the receiver end, and standard phase/amplitude measurements are conducted for INFT symbol categorization after the D2NN's signal conversion process. Moreover, the proposed D2NN demonstrates remarkable resilience against transmission system imperfections and phase resolution limitations.

Last but not the least, we'd like to comment on the feasibility of applying this method to INFT signals with a continuous spectrum. The proposed time domain

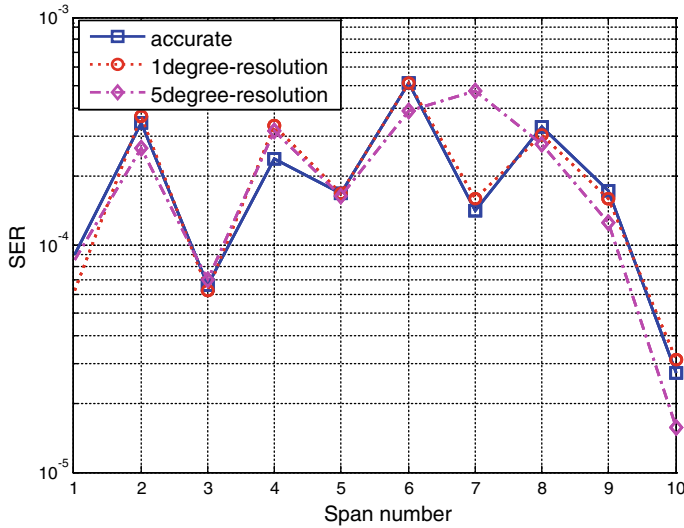


Fig. 7.17 The SER of the de-multiplexed signals with respect to fiber spans under the SNRs of 15 dB. The fiber span length is 80 km. Impact of the phase neuron resolution is shown in the figure. This figure is reprinted/adapted with permission from [Opt. Express 30, 38576–38586 (2022)] © The Optical Society of America [38]

D2NN functions as a unitary transformer, aiding in symbol recognition. For continuous spectra, performance may decline because NFT is not a unitary transform, and continuous spectrum signals cannot be fully converted by the time domain D2NN. Nevertheless, performance for both continuous and discrete spectrum recovery can be enhanced by incorporating a nonlinear activation function (via the Kerr effect) to make the transform non-unitary. Relevant research efforts are currently underway.

7.3.4 Summary

In conclusion, we have introduced a time domain D2NN implemented through a cascade of dispersion elements and phase modulators. We have also presented a BP algorithm that computes the phase gradient by inverse propagation of the conjugated error signal. Our newly developed D2NN has undergone testing as an all-optical INFT symbol recognizer, delivering impressive performance for a 1 GBaud Pol-Muxed NFT system. The proposed approach finds applications not only in NFT receivers but also in various other domains that demand all-optical time domain signal transformation and recognition, such as sensing, signal encoding and decoding, beam distortion compensation, and image recognition.

7.4 Spherical Gauss-Laguerre Beam Propagation in 4D Space-time

Part of the section reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62].

In the previous sections, it has been shown that the beam could propagate and being processed as if they were in space. In this section, the space and time are further united. A new category of beams, known as spherical Gauss-Laguerre beams, has been introduced. These beams exhibit stable propagation within anomalous dispersive media. Within this medium, the second-order derivative concerning time (t) can be integrated with the two-dimensional (2D) Laplacian operator in the transverse plane, forming a three-dimensional (3D) Laplacian operator. This operator describes the beam's propagation along the z -axis within the four-dimensional (4D) x - y - z - t space-time. By employing the variable separation method, we solve the wave equation and obtain the analytical expression for spherical Gauss-Laguerre beams. These beams exhibit a 3D Gaussian field distribution, featuring a beam waist that varies with propagation distance. Unlike traditional 2D spatial vortex beams, these 3D beams can possess either spatial vortices or spatiotemporal optical vortices (STOVs) depending on the choice of vortex plane within the 3D x - y - t space-time. The derived expression for spherical Gauss-Laguerre beams in 4D space-time has been validated through numerical simulations, showing excellent agreement.

7.4.1 Introduction

Beams are traditionally regarded as a spatial concept, confined within spatial coordinates. However, the recent integration of the time dimension into beam patterns has led to the emergence of various innovative beams, such as spatiotemporal optical vortex (STOV) beams, which exhibit vortices on a hybrid spatiotemporal plane [63–73].

Leading researchers have indicated that stable spatiotemporal beam propagation occurs in media with anomalous dispersion. In this context, the second-order derivative concerning time (t) can be merged with the second-order derivative related to spatial coordinates, forming the spatiotemporal Laplacian operator [74–77]. This approach has facilitated the discovery of several spatiotemporal beam classes in anomalous dispersive media, including the spatiotemporal 2D Bessel STOV beams [77], the 3D spatiotemporal localized pulse [74, 75], and the quasi-non-spreading spatiotemporal beams [76].

Pioneering researchers have investigated 2D beam propagation in anomalous dispersive media [77]. The principle to formulate beam propagation in x - z - t coordinates holds similarities to beam propagation in 3D spatial coordinates. Consequently, the beam pattern can be seamlessly transferred from spatial coordinates to combined spatiotemporal coordinates. This suggests that beam classes in x - z - t

coordinates resemble conventional spatial beam classes in x – y – z space, like Bessel beams, Gaussian beams, Hermite Gaussian beams, Laguerre Gaussian beams, and more.

Unlike the 3D scenario, spatiotemporal beam propagation in the 4D x – y – z – t spacetime relies on the 4D wave equation, resulting in beam patterns different from conventional spatial beams. For instance, S. Longhi proposed that a 3D diffraction-free beam in 4D spacetime resembles a sinc function shape, lacking any direct counterparts in 3D space [75]. Besides the sinc function-shaped beam, O. V. Borovkova et al. suggested that a broader class of quasi-non-spreading beams could emerge from a two-dimensional integral [76].

While numerous innovative spatiotemporal 3D beams have been proposed and examined in the 4D space–time [74–76], a significant challenge persists: diffraction-free spatiotemporal 3D beams, which have been derived using closed-form expressions, possess infinite power [75]. For instance, the sinc function-shaped beam, when integrated for power, yields an unbounded value, making it impractical for real-world applications. The approach for generating quasi-non-spreading beams may produce beams with finite power, but it lacks closed-form expressions [76]. Researchers are keen on discovering a closed-form Gaussian-shaped beam class in 4D space–time, which are similar to Hermite Gaussian and Laguerre Gaussian beams in 3D space, for easier practical generation.

In this study, we aim to derive a new class of closed-form Gaussian-shaped 3D beams in 4D space–time, called spherical Gauss-Laguerre beams (to distinguish them from cylindrical Laguerre Gaussian beams). Their expression comprises the product of the spherical harmonic function and the spherical Gauss-Laguerre function, ensuring beams with finite power. The spherical Gauss-Laguerre function concept was previously introduced and discussed in molecular science, serving as a basis function for solving the free-space Schrödinger equation [78]. Unlike the purely spatial spherical Gauss-Laguerre function described in [78], the spherical Gauss-Laguerre beams proposed here evolve with respect to z , featuring a z -dependent 3D beam waist. To the author's knowledge, this concept is novel in wave propagation and lacks precedents in traditional spatial beam research. The derived expression is validated through numerical simulations, showing excellent agreement.

7.4.2 Theory

Ignoring polarization-dependent effects, optical beam propagation in dispersive media can be described under the paraxial approximation by [77]:

$$\frac{\partial U}{\partial z} = \frac{1}{2jk} \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) + j \frac{\beta_2}{2} \frac{\partial^2 U}{\partial t^2}, \quad (7.33)$$

where j represents the imaginary unit, U denotes the beam envelope, k is the wave number within the media, x , y and z are Cartesian coordinates, and β_2 signifies the

dispersion coefficient. $t = t' - z/v_g$ reflects the converted time frame, where t' denotes the original time frame, and v_g stands for the group velocity.

In anomalous dispersion media, where $\beta_2 < 0$, the second-order derivative concerning t can be integrated into the 2D Laplacian operation, forming a 3D Laplacian operator [77]:

$$\frac{\partial U}{\partial z} = \frac{1}{2jk} \left(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial T^2} \right), \quad (7.34)$$

where $T = t/\sqrt{|\beta_2|k}$ is the normalized time, which has the unit of length.

It is possible to convert the Cartesian coordinates (x, y, T) into the spherical coordinates (r, θ, φ) as is illustrated in Fig. 7.18.

And derive the spherical coordinates equation:

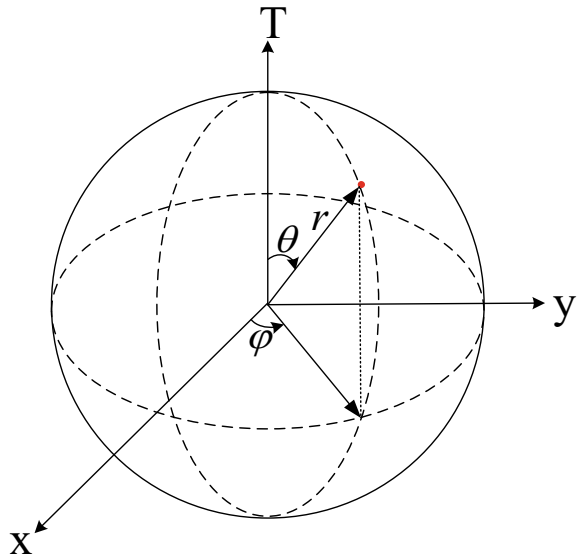
$$\frac{\partial U}{\partial z} = \frac{1}{2jk} \left(\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right) U. \quad (7.35)$$

To solve Eq. (7.35), we can postulate:

$$U = R(z, r) \Psi(\theta, \varphi) \quad (7.36)$$

Combining Eqs. (7.36) and (7.35) and applying the variable separation method, one can deduce the equations for R and Ψ :

Fig. 7.18 The (r, θ, φ) coordinates. This figure is reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62]



$$\begin{cases} \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) \Psi + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \Psi + \lambda \Psi = 0, \\ \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) R + \left(-2jk \frac{\partial R}{\partial z} - \frac{\lambda}{r^2} R \right) = 0. \end{cases} \quad (7.37)$$

where λ is a constant to be evaluated. The first equation concerning $\Psi(\theta, \varphi)$ is solvable and its solution is the spherical harmonic function, which can be obtained by decompose $\Psi(\theta, \varphi)$ as:

$$\Psi = \Theta(\theta) \Phi(\varphi). \quad (7.38)$$

We may have:

$$\begin{cases} \frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d}{d\theta} \right) \Theta - \frac{m^2}{\sin^2 \theta} \Theta + \lambda \Psi = 0, \\ \frac{d^2 \Phi}{d\varphi^2} + m^2 \Phi = 0, \end{cases} \quad (7.39)$$

where m^2 is a constant. Considering $\Phi(\varphi)$ is periodical with respect to φ , m ought to be an integer. Additionally, to guarantee the convergence of Θ , we set $\lambda = l(l+1)$, where l is a non-negative integer. The integer m can range from $-l$ to l . Equation (7.39) leads to the well-known spherical harmonic function as the solution for Ψ :

$$\Psi \propto Y_{lm}(\theta, \varphi) = P_l^m(\cos \theta) e^{im\varphi}, \quad (7.40)$$

where P_l^m is the Legendre function with order l and degree m , with m also denoting the vortex or the beam topological charge.

Using the computed constant $\lambda = l(l+1)$, we formulate the radial equation for the function R as

$$\frac{\partial^2 R}{\partial r^2} + \frac{2}{r} \frac{\partial R}{\partial r} + \left(-2jk \frac{\partial R}{\partial z} - \frac{l(l+1)}{r^2} R \right) = 0. \quad (7.41)$$

Though the transformation $R = ur^{-1/2}$, one has

$$\frac{\partial^2}{\partial r^2} u + \frac{1}{r} \frac{\partial}{\partial r} u + \left(-2jk \frac{\partial}{\partial z} - \frac{(l+1/2)^2}{r^2} \right) u = 0. \quad (7.42)$$

Equation (7.42) is similar to the 2D radial equation in the polar coordinates:

$$\frac{\partial^2}{\partial \rho^2} v + \frac{1}{\rho} \frac{\partial}{\partial \rho} v + \left(-2jk \frac{\partial}{\partial z} - \frac{m^2}{\rho^2} \right) v = 0, \quad (7.43)$$

where v represents the radial function of the 2D beams in the spatial cylindrical coordinates, and ρ denotes the polar coordinates radius. It is acknowledged that the

Laguerre Gaussian function serves as the solution to Eq. (7.43), providing the radial distribution function for the 2D spatial Laguerre Gaussian beam:

$$\begin{aligned}
 v_{pn}(\rho, z) &\propto \left(\frac{1}{w(\zeta)}\right)^2 \left(\frac{\sqrt{2}\rho}{w(\zeta)}\right)^n L_p^n\left(\frac{2\rho^2}{w^2(\zeta)}\right) \exp\left(-\frac{\rho^2}{w_0^2(1-j\zeta)} + j\Psi_{pn}\right), \\
 \rho^2 &= x^2 + y^2, \\
 \varphi &= \tan^{-1}\left(\frac{y}{x}\right), \\
 z_R &= \frac{kw_0^2}{2}, \\
 \zeta &= \frac{z}{z_R}, \\
 w(\zeta) &= w_0\sqrt{1+\zeta^2}, \\
 \Psi_{pn} &= (1+n+2p)\xi, \\
 \xi &= \tan^{-1}(\zeta),
 \end{aligned} \tag{7.44}$$

where w_0 denotes the initial beam waist, L_p^n denotes the generalized Laguerre polynomial with radial order p and azimuth order n , ζ denotes the reduced length, $w(\zeta)$ denotes the beam waist at the corresponding reduced length, and ξ denotes the Gouy phase. Comparing Eq. (7.42) with Eq. (7.43), we observe that the only difference lies in the replacement of m with $l + 1/2$. Hence, by replacing m with $l + 1/2$ in Eq. (7.44), one has:

$$\begin{aligned}
 u &\propto \left(\frac{1}{w(\zeta)}\right)^2 \left(\frac{\sqrt{2}r}{w(\zeta)}\right)^{l+1/2} L_p^{l+1/2}\left(\frac{2r^2}{w^2(\zeta)}\right) \\
 &\exp\left(-\frac{r^2}{w_0^2 - 2jz/k}\right) \exp\left(j\left(2p + l + \frac{3}{2}\right)\xi\right).
 \end{aligned} \tag{7.45}$$

Combining the solutions for R and Ψ , one arrives at:

$$\begin{aligned}
 U &= R(z, r)\Psi(\theta, \varphi) = r^{-1/2}u\Psi(\theta, \varphi) \\
 &\propto \left(\frac{1}{w(\zeta)}\right)^{3/2} \left(\frac{\sqrt{2}r}{w(\zeta)}\right)^l L_p^{l+1/2}\left(\frac{2r^2}{w^2(\zeta)}\right) Y_{lm}(\theta, \varphi) \\
 &\exp\left(-\frac{r^2}{w_0^2 - 2jz/k}\right) \exp\left(j\left(2p + l + \frac{3}{2}\right)\xi\right).
 \end{aligned} \tag{7.46}$$

By multiplying a factor of $w_0^{3/2}$, we ensure that the lowest-order beam has a unit peak value, resulting in Eq. (7.47). This equation describes the novel 3D spherical Gauss-Laguerre beam propagation in 4D space-time.

$$\begin{aligned}
U &\propto \left(\frac{w_0}{w(\zeta)}\right)^{3/2} \left(\frac{\sqrt{2}r}{w(\zeta)}\right)^l L_p^{l+1/2}\left(\frac{2r^2}{w^2(\zeta)}\right) Y_{lm}(\theta, \varphi) \\
&\quad \exp\left(-\frac{r^2}{w_0^2 - 2jz/k}\right) \exp\left(j\left(2p + l + \frac{3}{2}\right)\xi\right), \\
Y_{lm}(\theta, \varphi) &= P_l^m(\cos \theta) e^{im\varphi},
\end{aligned} \tag{7.47}$$

where w_0 represents the initial beam waist, l denotes a non-negative integer, and m signifies the vortex of the beam. m is an integer that can range from $-l$ to l , indicating that for each value of l , there can be $(2l + 1)$ possible m values, spanning from $-l$ to $+l$, P_l^m represents the associated Legendre function with order l and degree m , $L_p^{l+1/2}$ stands for the generalized Laguerre polynomial, featuring a radial order of p and a azimuth order of $l + 1/2$, ζ is known as the reduced length, $w(\zeta)$ refers to the beam waist at the reduced length, and ξ signifies the Guoy phase which are:

$$\begin{aligned}
\zeta &= \frac{z}{z_R}, \\
z_R &= \frac{kw_0^2}{2}, \\
w(\zeta) &= w_0 \sqrt{1 + \zeta^2}, \\
\xi &= \tan^{-1}(\zeta).
\end{aligned} \tag{7.48}$$

Several intriguing properties can be deduced from Eq. (7.47). The radial distribution of the beam shape is not influenced by the value of m , but rather by l . This implies that the beam structure remains unaffected as the topological charge increases, unlike conventional spatial vortex beams that expand. Secondly, the vortex exists on the φ plane, which corresponds to the x - y plane in Fig. 7.18. However, one could alternatively choose the x - T plane or y - T plane as the φ plane. This flexibility suggests that the beams are capable of sustaining a vortex on both the spatial and spatiotemporal planes, making Eq. (7.47) a generalized expression applicable to both spatial orbital angular momentum (OAM) beams and spatiotemporal optical vortex (STOV) beams. Additionally, the plane can be any plane containing the origin, introducing a new category of space-time interactive beams.

7.4.3 Results and Discussions

Numerical simulations have been carried out to validate the accuracy of the equation. The beam was discretized on a 3D grid of $200 \times 200 \times 200$ with dimensions of $100 \mu\text{m} \times 100 \mu\text{m} \times 100 \mu\text{m}$. The numerical simulation of Eq. (7.34) relies on the Fourier transform. The beam is presumed to have a wavelength of 1550 nm and an initial beam width of $10 \mu\text{m}$. The propagation distance is set at $500 \mu\text{m}$,

which corresponds to 2.5 times the Rayleigh length. For simplicity and without compromising generality, the vortex plane of the beam is selected as the x - y plane, which can be effortlessly transformed into the x - T plane to generate a beam with spatiotemporal optical vortex (STOV).

Initially, we set the beam parameters as $l = 1$, $m = 0$, and $p = 0$. Given the 3D nature of the beam, we illustrate the 2D amplitude and phase distributions of the beam at $T = 0 \mu\text{m}$ in Fig. 7.19. By comparing the simulation results with those derived from the analytical formula, namely Eq. (7.47), it can be inferred that the formula is precise, and the recently discovered spherical Gauss-Laguerre beams indeed exist in the space-time.

In Fig. 7.20, the beam parameters are adjusted to $l = 1$, $m = 1$, and $p = 0$ with a spiral phase pattern carrying a topological charge of 1. The accordance between the numerical simulations and analytical outcomes provides additional confirmation of the accuracy of the spherical Gauss-Laguerre beam formula. Upon analysis of

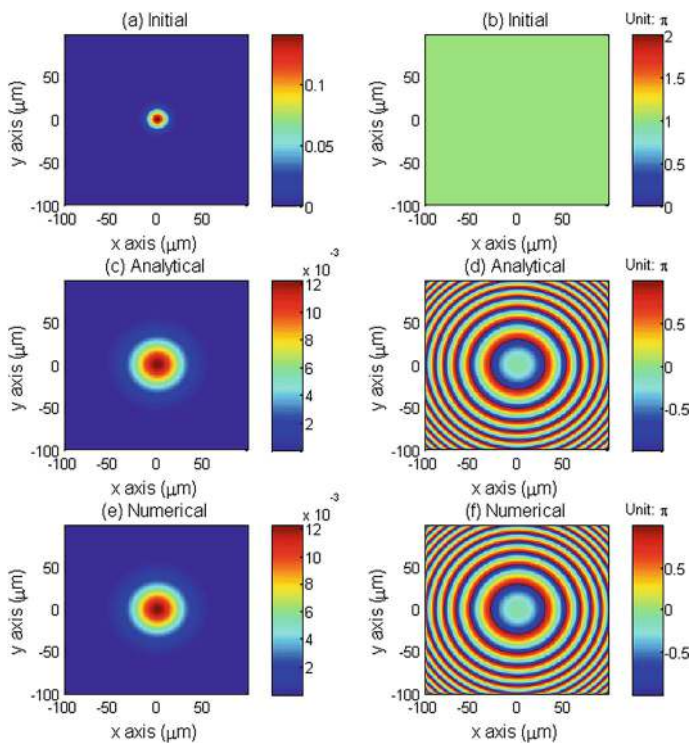


Fig. 7.19 Amplitude and phase distributions of the spherical Gauss-Laguerre beam ($l = 1$, $m = 0$, and $p = 0$) propagates for $500 \mu\text{m}$ in the z direction when $T = 0 \mu\text{m}$. **a** The amplitude of the initial beam. **b** The phase of the initial beam. **c** The amplitude by the analytical formula. **d** The phase by the analytical formula. **e** The amplitude by the numerical simulations. **f** The phase by the numerical simulations. This figure is reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62]

Figs. 7.19 and 7.20, it becomes evident that the beam radius experiences much expansion as the topological charge progresses from 0 to 1, distinguishing it from conventional spatial vortex beams. Notably, the beam radius is influenced by the parameter ' l ' and not ' m '; similarly, the Guoy phase of the beam follows the same rule.

Lastly, in Fig. 7.21, the beam parameters are modified to $l = 1$, $m = 1$, and $p = 1$, signifying an increase in the beam's radial order from 0 to 1, leading to the formation of multiple rings within the beam. As is illustrated in the figure, the beam now exhibits two rings, both carrying a topological charge of 1. Once again, the numerical findings align precisely with the analytical predictions.

The sliced output of the beam, corresponding to various values of T , is presented in Fig. 7.22. This figure illustrates the wave packet distribution of the beam depicted in Fig. 7.4 in a three-dimensional format. Evidently from Fig. 7.22, it is apparent that the beam follows a Gaussian distribution in the T frame as T ranges from 5 to

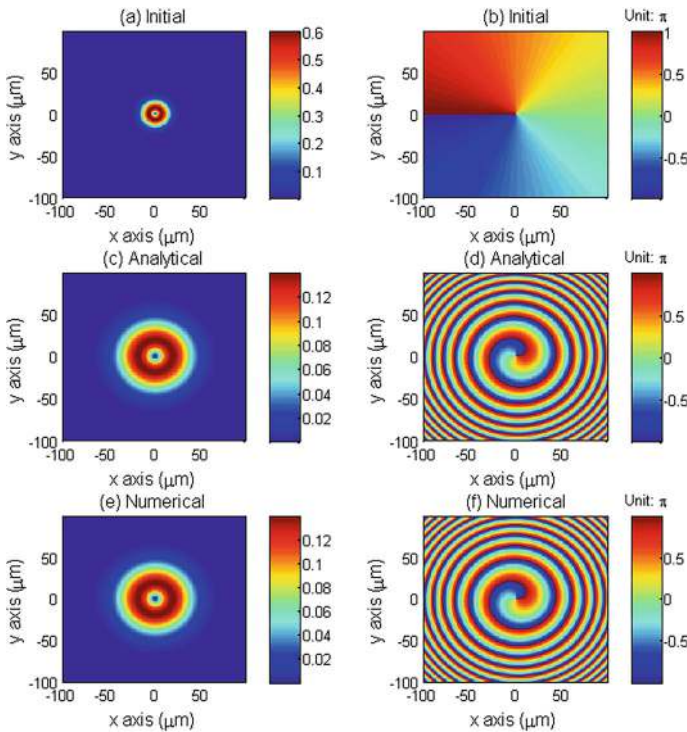


Fig. 7.20 Amplitude and phase distributions of the spherical Gauss-Laguerre beam ($l = 1$, $m = 1$, and $p = 0$) propagates for $500 \mu\text{m}$ in the z direction when $T = 0 \mu\text{m}$. **a** The amplitude of the initial beam. **b** The phase of the initial beam. **c** The amplitude by the analytical formula. **d** The phase by the analytical formula. **e** The amplitude by the numerical simulations. **f** The phase by the numerical simulations. This figure is reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62]

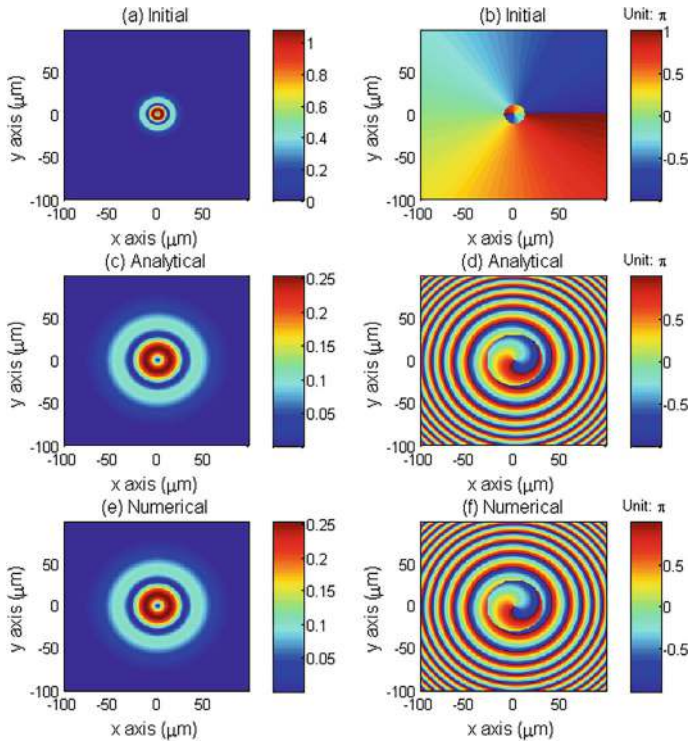


Fig. 7.21 Amplitude and phase distributions of the spherical Gauss-Laguerre beam ($l = 1$, $m = 1$, and $p = 1$) propagates for $500 \mu\text{m}$ in the z direction when $T = 0 \mu\text{m}$. **a** The amplitude of the initial beam. **b** The phase of the initial beam. **c** The amplitude by the analytical formula. **d** The phase by the analytical formula. **e** The amplitude by the numerical simulations. **f** The phase by the numerical simulations. This figure is reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62]

$80 \mu\text{m}$. It's worth noting that the numerical and analytical results align remarkably well in this instance too.

To sum up, the proposed spherical Gauss-Laguerre beams indeed exist. The beam's propagation mechanism differs significantly from the conventional Laguerre Gaussian beam, and its field distribution evolves in four-dimensional space-time, adhering to the derived formula, namely Eq. (7.47). Topological charges can manifest either in space or space-time.

To generate such a beam, one could utilize an arbitrary spatiotemporal field generator, as is mentioned in [79]. This generator comprises a spectral pulse shaper and a multi-plane light conversion (MPLC) device. During experimentation, it is crucial to ensure that the normalized time frame is appropriately scaled with the two spatial coordinates, namely x and y .

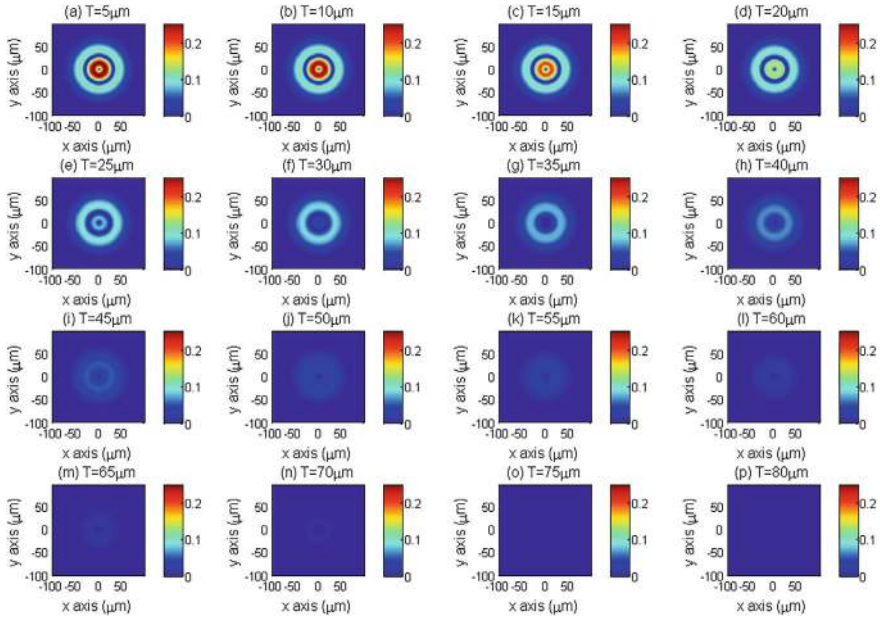


Fig. 7.22 Amplitude distributions of the spherical Gauss-Laguerre beam ($l = 1$, $m = 1$, and $p = 1$) propagates for $500 \mu\text{m}$ in the z direction when T have different values. **a–p** $T = 5$ to $80 \mu\text{m}$. This figure is reprinted/adapted with permission from [Opt. Express 23, 38334–38342 (2023)] © The Optical Society of America [62]

7.4.4 Summary

In conclusion, we have introduced a unique category of 3D spatiotemporal beams that propagate within the 4D x – y – z – t space–time, featuring a Gaussian field distribution. These beams can exhibit either spatial vortices or STOVs, depending on the chosen vortex plane, distinguishing them from any existing beams in spatial coordinates. Additionally, unlike spatial vortex beams, whose beam diameters and Guoy phases increase linearly with the topological charge, our proposed spherical Gauss-Laguerre beams maintain constant beam diameters and Guoy phases, independent of the topological charge. The precise mathematical expression for these newly discovered beams has been validated through numerical simulations, demonstrating an excellent match.

Given that these beams operate within the 4D x – y – z – t space–time, they differ significantly from traditional space–time beams. This suggests that they can provide an additional degree of freedom for light manipulation. For instance, these beams can be simultaneously multiplexed on both spatial vortices and STOVs, potentially enhancing information carrying capacity in free-space optical communication systems.

7.5 Spatiotemporal Diffractive Deep Neural Networks

Part of the section reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80].

With the combination of the space and time, one may have the new beam and a new tool is needed to process it. In this section, we propose a spatiotemporal diffractive deep neural network (STD2NN) tailored for spatiotemporal signal processing. The STD2NN consists of gratings that transform signals from the frequency domain to the spatial domain, along with multiple layers made up of spatial lenses and space light modulators (SLMs) responsible for spatiotemporal phase modulation. We introduce an all-optical backpropagation (BP) algorithm for fine-tuning the SLM phases. The gradient of the loss function is determined by computing the inner product of the forward-propagating optical field and the backward-propagating conjugated error field. To demonstrate our concept, we use the STD2NN to generate a spatiotemporal word, “OPTICA.” Furthermore, we demonstrate a spatiotemporal optical vortex (STOV) beam multiplexer based on the STD2NN. This multiplexer converts spatially separated Gaussian beams into STOV wave-packets carrying distinct topological charges. Both examples highlight the STD2NN’s proficiency in generating and manipulating spatiotemporal signals.

7.5.1 Introduction

The combination of the spatial and temporal concepts gives rise to novel physical phenomena. For instance, a spatiotemporal localized wave packet, analogous to the non-diffractive Bessel beam in space, can persist during propagation [81]. Such waves deviate from pure spatial beams or temporal pulses due to the inseparability of their spatial and temporal spectra [82]. The other recent revelation in spatiotemporal optics is the spatiotemporal optical vortex (STOV), which extends the purely spatial concept of the optical vortex into the four-dimensional space–time [63–67, 83, 84]. This concept has ushered in a new era of structured light research.

As research in this field flourishes, bringing numerous discoveries, tools for complex spatiotemporal signal processing have gained prominence. It is well-known that gratings can transform signals from the frequency domain to the spatial domain, followed by spatiotemporal modulation using a spatial light modulator (SLM) [83, 85, 86]. H. E. Kondakci et al. proposed generating spatiotemporal diffraction-free beams using a grating and an SLM [86], while S. Wang et al. suggested creating STOV wave-packets with a similar experimental setup [83]. M. Mounaix et al. argued that multi-plane light conversion (MPLC) offered more robust and precise shaping of spatiotemporal waves [79], albeit requiring a relatively intricate spectral pulse shaper and 14 phase planes with phases optimized via simulated annealing. Alternatively, H. Wang et al. introduced a customized photonic crystal slab with unique spatiotemporal dispersive properties for STOV wave-packet generation [84]. L. Chen

et al. proposed using metasurfaces to shape spatiotemporal waves [87]. Despite these groundbreaking studies [63–67, 79, 83–87], there remains a significant need for a versatile tool for spatiotemporal signal processing. Such a tool should generate arbitrary spatiotemporal signals or multiplex/demultiplex different spatiotemporal symbols, such as STOV wave-packets with varying topological charges, while maintaining low complexity.

However, it requires either temporal integration at the detector or temporal buffers, limiting its application to low-speed (in the order of nanoseconds [90]) spatiotemporal signal processing. This makes it unsuitable for high-speed spatiotemporal wave-packets, such as STOV wave-packets (in the order of 0.1 ps).

Neural networks, as robust signal processing tools, have undergone significant advancements with diverse applications [50]. Recently, an all-optical neural network, known as the diffractive deep neural network (D2NN), has surfaced as a notable milestone in the field [51–58, 88–90]. This technology enables light-speed, all-optical signal processing and recognition with exceptionally low power consumption. Unlike conventional neural networks that feature nonlinear activation components, the D2NN accomplishes signal processing through layered diffraction. Typically, the D2NN concept is applied in the spatial domain, utilizing cascading free space propagation and phase-tuning elements for spatial optical wave transformation. We introduced a time domain D2NN that employs dispersive elements to simulate “free propagation” in the time domain, followed by time domain phase modulators [88]. Recently, a D2NN tailored for spatiotemporal signal processing has been proposed [89, 90]. However, it requires either temporal integration at the detector or temporal buffers, limiting its application to low-speed (in the order of nanoseconds [90]) spatiotemporal signal processing. This makes it unsuitable for high-speed spatiotemporal wave-packets, such as STOV wave-packets (in the order of 0.1 ps).

In this section, we introduce a novel architecture for spatiotemporal signal processing, known as the spatiotemporal diffractive deep neural network (STD2NN). The STD2NN consists of input and output gratings, as well as multiple layers, where each layer comprises a 4f system with a SLM integrated within. Initially, the signal reflects from an input grating, which converts the frequency domain signal to an x-dimensional signal in the spatial domain. Consequently, the signal carries frequency information along the x-dimension and spatial information along the y-dimension. It undergoes distinct spatial and temporal propagation within the two dimensions of the 4f system, with phase adjustments facilitated by the SLM. To efficiently calibrate the SLM phases, training is conducted using an all-optical backpropagation (BP) algorithm. This algorithm computes the phase gradient using both the forward propagating field and the backward propagating conjugated error field, significantly enhancing the optimization process and facilitating real-time phase adjustments. At the output stage, the signal reflects from the output grating, which results in the generation of the desired spatiotemporal signal. As a demonstration of our concept, we have generated a spatiotemporal word “OPTICA” using the STD2NN. Furthermore, we illustrate a STOV wave-packet multiplexer based on the STD2NN, which transforms spatially separated Gaussian beams into STOV wave-packets possessing varying topological charges. Both examples highlight the STD2NN’s versatility in

generating and processing arbitrary spatiotemporal signals, indicating its potential for comprehensive pulse shaping in the space–time domain simultaneously.

7.5.2 The STD2NN Architecture and All Optical Error Back-Propagation

The proposed STD2NN consists of two gratings and multiple neuron layers, where each layer is formed by a $4f$ pulse shaper with an embedded SLM in the middle. The basic structure of the STD2NN can be seen in Fig. 7.23.

At the input of the STD2NN, the signal $S_0(x_0, y_0, \omega)$ with the spatial variables x_0, y_0 and the frequency variable ω undergoes reflection from a grating, transforming it into U_0 , which serves as the input for the $4f$ system [83, 85]:

$$U_0(x_0, y_0, \omega) = \sqrt{\beta} S_0(\beta x_0, y_0, \omega) \exp(j\gamma \omega x_0), \quad (7.49)$$

where $\beta = \cos(\theta_i)/\cos(\theta_d)$ with θ_i and θ_d as the incident angle and the diffraction angle, $\gamma = k/(d\cos(\theta_d))$ with k as the free space wavenumber and d as the grating line spacing. The inclusion of diffractive gratings at the input and output facilitates a unique mapping of specific frequency components to precise points on the x-axis, and this converts the conventional spatial D2NN into the spatiotemporal D2NN.

The $4f$ pulse shaper consists of two cylindrical lenses [83, 85] which focus the beams in the x direction. While in the y -direction, beams undergo free space propagation.

Neglecting the x dimension, beam propagation in the y -dimension over a distance of $2f$ follows the principles of free space propagation:

$$U'_1(y, \omega) = IFT_y \left(\exp \left(-\frac{jk_y^2 f}{k} \right) FT_y(U_0(y, \omega)) \right), \quad (7.50)$$

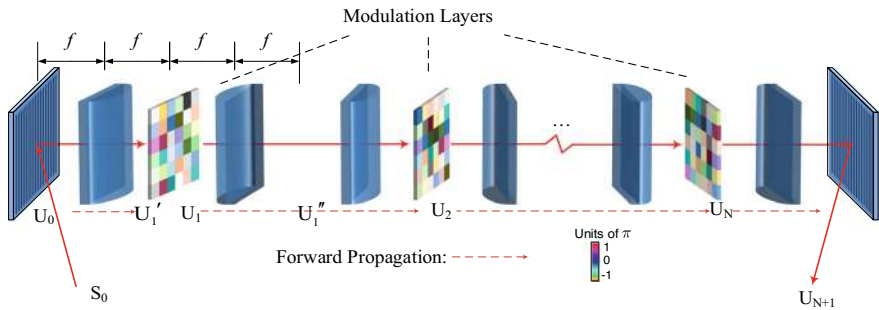


Fig. 7.23 Basic structure of the STD2NN. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

where FT_y and IFT_y denote the y dimensional Fourier transform and inverse Fourier transform. Neglecting the y dimension, the beam in the x -dimension experiences the Fourier transform induced by the lens:

$$\begin{aligned} U'_1(x, \omega) &= \frac{1}{\sqrt{\lambda f}} FT_x(U_0)\left(\frac{kx}{f}, \omega\right) \\ &= \frac{1}{\sqrt{\lambda f}} FT_x\left(\sqrt{\beta} S_0(\beta x, \omega) \exp(j\gamma \omega x)\right)\left(\frac{kx}{f}, \omega\right) \\ &= \frac{1}{\sqrt{\beta \lambda f}} FT_x(S_0)\left(\frac{1}{\beta} \left(\frac{kx}{f} - \gamma \omega\right), \omega\right), \end{aligned} \quad (7.51)$$

where FT_x denotes the Fourier transform in the x dimension. Combining Eq. (7.50 and 7.51), one arrives at Eq. (7.52). Neglecting the phase term $\exp(2jkf)$, the field U'_1 on the focal plane of the first lens, prior to spatiotemporal modulation, is given by [83, 85]:

$$\begin{aligned} \tilde{U}'_1(x_1, k_y, \omega) &= \frac{1}{\sqrt{\lambda f}} \exp\left(-\frac{jk_y^2 f}{k}\right) \tilde{U}_0\left(\frac{kx_1}{f}, k_y, \omega\right) \\ &= \frac{1}{\sqrt{\lambda f \beta}} \exp\left(-\frac{jk_y^2 f}{k}\right) \tilde{S}_0\left(\frac{1}{\beta} \left(\frac{kx_1}{f} - \gamma \omega\right), k_y, \omega\right), \end{aligned} \quad (7.52)$$

where f denotes the lens focal length, λ the free space wavelength, and k_y the y directional spatial frequency. $\tilde{U}$ denotes the y dimensional Fourier transform, and $\tilde{\tilde{U}}$ denotes the two-dimensional Fourier transform in the x - y dimensions.

If the incident beam field distribution follows the x directional plane wave, e.g., it keeps a constant amplitude 1 in the x direction, one arrives at:

$$\tilde{U}'_1(x_1, k_y, \omega) = \frac{2\pi}{\sqrt{\lambda f \beta}} \exp\left(-\frac{jk_y^2 f}{k}\right) \delta\left(\frac{1}{\beta} \left(\frac{kx_1}{f} - \gamma \omega\right)\right) \tilde{S}_0(k_y, \omega), \quad (7.53)$$

where δ represents the Dirac function. A direct correlation between the x dimension in the spatial domain and the frequency domain ω is found on the Fourier plane. It is worth noting that when the beam width in the x direction is sufficiently large, it can be approximated as a plane wave.

Hence, we can employ an SLM for spatiotemporal modulation, resulting in a field U_1 after the SLM:

$$U_1(x_1, y_1, \omega) = U'_1(x_1, y_1, \omega) M(x_1, y_1), \quad (7.54)$$

where M denotes the spatiotemporal phase modulation mask induced by the SLM, which can be considered as the controllable neurons on the first layer of the STD2NN.

This modulation mask consists of multiple cells organized in a two-dimensional array, with each cell functioning as an individual neuron. The phase of each neuron can be adjusted between 0 and 2π . This neuron array facilitates spatial modulation in the y dimension and temporal modulation in the x dimension, as is mentioned earlier. Distinct from conventional neural networks, diffractive deep neural networks eliminate the need for weight matrices.

As the field propagates through a second lens with the same focal length f , we obtain the field U_1'' on the focal plane as:

$$\tilde{U}_1''(x_2, k_y, \omega) = \frac{1}{\sqrt{\lambda f \beta}} \exp\left(-\frac{jk_y^2 f}{k}\right) \tilde{U}_1\left(\frac{kx_2}{f}, k_y, \omega\right), \quad (7.55)$$

In the space-frequency domain, one has:

$$\tilde{\tilde{U}}_1''(k_x, k_y, \omega) = \sqrt{\frac{\lambda f}{\beta}} \exp\left(-\frac{jk_y^2 f}{k}\right) \tilde{U}_1\left(-\frac{fk_x}{k}, k_y, \omega\right). \quad (7.56)$$

The aforementioned equations constitute a detailed mathematical model for beam evolution within the first layer of the STD2NN. For intermediate layers, a $4f$ system is still employed, and the field evolves according to Eqs. (7.52–7.56) in a cascading manner. This configuration ensures that different frequency components are mapped to distinct points on the SLM's x -axis. A fully diffractive approach could lead to mixed frequency components at a single SLM point, making f - y dimensional modulation impractical.

While this model is suitable for numerical calculations of spatiotemporal signals, the SLM phase is typically discretized. Therefore, a method for adjusting these phases to generate the desired spatiotemporal signal is necessary. To this end, we introduce a discretized model which helps to derive the BP algorithm for SLM phase tuning.

Based on our analysis, if the beam satisfies the plane wave approximation in the x direction, it can be regarded as a two-dimensional t - y (or ω - y) spatiotemporal (or spatial and frequential) symbol on the SLM plane as it transmits through the STD2NN. Consequently, we can discretize the system in both space and frequency, and represent the optical field after each layer's SLM as a matrix. In this matrix, rows correspond to the frequency domain numbering, while columns represent the spatial domain numbering. This matrix is related to the optical field after the SLM of the preceding layer:

$$\mathbf{U}_k = \left(\mathbf{D}_k^s \mathbf{U}_{k-1} \mathbf{D}_k^f\right) \odot \mathbf{M}_k, \quad (7.57)$$

where matrix $\mathbf{U}_k$ stands for the optical field after the SLM of the k th layer of the ST-D2NN, $\mathbf{D}_k^f$ and $\mathbf{D}_k^s$ are the frequency domain and spatial domain propagation operators for the x dimensional lenses and the y dimensional free space propagation, $\mathbf{M}_k$ is a phase operator introduced by the SLM, and $\odot$ denotes element-wise product.

While $\mathbf{U}_k$ is a matrix, $\mathbf{D}_k^f$ and $\mathbf{D}_k^s$ are two square matrices with their sizes equaling the number of rows and the number of columns in $\mathbf{U}_k$. $\mathbf{M}_k$ is a matrix with the same size as $\mathbf{U}_k$, and the element of $\mathbf{M}_k$ on the l th row and the m th column is denoted as $n_{k,lm}$, which is considered as the neuron of the STD2NN, with $n_{k,lm} = \exp(j\phi_{k,lm})$. Assuming that the STD2NN consists of N layers of neurons for, one has:

$$\mathbf{U}_N = \left(\mathbf{D}_N^s \cdots \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \mathbf{M}_k \cdots \mathbf{D}_N^f \right) \odot \mathbf{M}_N, \quad (7.58)$$

where $\mathbf{U}_0$ stands for the input signal and $\mathbf{U}_N$ stands for the signal after the SLM of the output layer. At last, one has:

$$\mathbf{U}_{N+1} = \mathbf{D}_{N+1}^s \mathbf{U}_N \mathbf{D}_{N+1}^f, \quad (7.59)$$

where $\mathbf{U}_{N+1}$ represents the output beam after reflection from the second grating, which possesses the same physical parameters as the input grating but is positioned symmetrically with respect to it, $\mathbf{D}_{N+1}^f$ and $\mathbf{D}_{N+1}^s$ denote the propagation operators in the frequency and spatial domains, respectively. In order to tune the neuron phases, one requires the BP algorithm, which is derived by defining the loss function at first:

$$L = |\mathbf{U}_{N+1} - \mathbf{T}|^2 = \text{Trace}((\mathbf{U}_{N+1} - \mathbf{T})^H (\mathbf{U}_{N+1} - \mathbf{T})), \quad (7.60)$$

where $\mathbf{T}$ denotes the target wave. The derivation of the BP algorithm will be elaborated later. Before that, it's crucial to determine the gradient of the field concerning the neuron:

$$\begin{aligned} \frac{\partial \mathbf{U}_{N+1}}{\partial n_{k,lm}} &= \mathbf{D}_{N+1}^s \left(\left(\mathbf{D}_N^s \cdots \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \frac{\partial \mathbf{M}_k^s}{\partial n_{k,lm}} \cdots \mathbf{D}_N^f \right) \odot \mathbf{M}_N \right) \mathbf{D}_{N+1}^f \\ &= \mathbf{D}_{N+1}^s \left(\left(\mathbf{D}_N^s \cdots \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \mathbf{1}_{lm} \cdots \mathbf{D}_N^f \right) \odot \mathbf{M}_N \right) \mathbf{D}_{N+1}^f. \end{aligned} \quad (7.61)$$

$$\mathbf{1}^{lm}(i, j) = \begin{cases} 1 & i = l \text{ and } j = m \\ 0 & \text{otherwise} \end{cases}$$

One may compute the gradient of the L function with respect to the neuron as:

$$\begin{aligned}
\frac{\partial L}{\partial n_{k,lm}} &= \text{Trace} \left((\mathbf{U}_{N+1} - \mathbf{T})^H \frac{\partial \mathbf{U}_{N+1}}{\partial n_{k,lm}} \right) \\
&= \text{Trace} \left((\mathbf{U}_{N+1} - \mathbf{T})^H \left(\mathbf{D}_{N+1}^s \left(\mathbf{D}_N^s \cdots \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \mathbf{1}^{lm} \cdots \mathbf{D}_N^f \right) \right) \right. \\
&\quad \left. \odot \mathbf{M}_N \mathbf{D}_{N+1}^f \right) \\
&= \text{Trace} \left(\mathbf{D}_{k+1}^f \left(\cdots \mathbf{D}_N^f \left(\mathbf{D}_{N+1}^f (\mathbf{U}_{N+1} - \mathbf{T})^H \mathbf{D}_{N+1}^s \right) \odot \mathbf{M}_N^T \mathbf{D}_N^s \cdots \right) \right. \\
&\quad \left. \odot \mathbf{M}_{k+1}^T \mathbf{D}_{k+1}^s \left(\left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \mathbf{1}^{lm} \right) \right) \\
&= \mathbf{P}_k^b(l, m) \mathbf{Q}_{klm}^f(l, m), \\
\mathbf{Q}_k^f &= \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right), \\
\mathbf{P}_k^b &= \left(\mathbf{D}_{k+1}^{sT} \left(\cdots \mathbf{D}_N^{sT} \left(\mathbf{D}_{N+1}^{sT} (\mathbf{U}_{N+1} - \mathbf{T})^* \mathbf{D}_{N+1}^{fT} \right) \odot \mathbf{M}_N^T \mathbf{D}_N^{fT} \cdots \right) \odot \mathbf{M}_{k+1}^T \mathbf{D}_{k+1}^{fT} \right).
\end{aligned} \tag{7.62}$$

Here, $n_{k,lm}$ and its complex conjugate are treated as independent variables. Notably, the derivative of $\mathbf{U}_{N+1}^H$ with respect to $n_{k,lm}$ is zero. Equation (7.62) implies that the derivative concerning the neuron located on the l th row and m th column relates to the product of corresponding elements from matrices $\mathbf{Q}_k^f$ and $\mathbf{P}_k^b$. Consequently, we can express this relationship in matrix notation as:

$$\frac{\partial L}{\partial \mathbf{n}_k} = \mathbf{P}_k^b \odot \mathbf{Q}_k^f, \tag{7.63}$$

Given that the neuron is tunable in phase, one has

$$\begin{aligned}
\frac{\partial L}{\partial \varphi_k} &= j \frac{\partial L}{\partial n_{k,lm}} n_{k,lm} + c.c. \\
&= 2\text{Re} \left(j \frac{\partial L}{\partial n_{k,lm}} n_{k,lm} \right)
\end{aligned} \tag{7.64}$$

where c.c. stands for complex conjugate. One further formulates Eq. (7.65) as the gradient of the phase matrix φ_k with respect to the loss function:

$$\begin{aligned}
\frac{\partial L}{\partial \varphi_k} &= 2\text{Re} \left(j \mathbf{P}_k^b \odot \mathbf{P}_k^f \right), \\
\mathbf{P}_k^f &= \left(\mathbf{D}_k^s (\cdots \mathbf{U}_0 \cdots) \mathbf{D}_k^f \right) \odot \mathbf{M}_k, \\
\mathbf{P}_k^b &= \left(\mathbf{D}_{k+1}^{sT} \left(\cdots \mathbf{D}_N^{sT} \left(\mathbf{D}_{N+1}^{sT} (\mathbf{U}_{N+1} - \mathbf{T})^* \mathbf{D}_{N+1}^{fT} \right) \odot \mathbf{M}_N^T \mathbf{D}_N^{fT} \cdots \right) \odot \mathbf{M}_{k+1}^T \mathbf{D}_{k+1}^{fT} \right),
\end{aligned} \tag{7.65}$$

where ^{T} denotes matrix transpose. The field $\mathbf{P}_k^f$ can be understood as the forward propagating optical field after the SLM of the k th layer, while the field $\mathbf{P}_k^b$ can be viewed as the backward propagating field after the SLM of the k th layer, starting the backward propagation from the output of the STD2NN with the conjugated field $(\mathbf{U}_{N+1} - \mathbf{T})^*$. The reciprocal property of the optical element [91] suggests that when

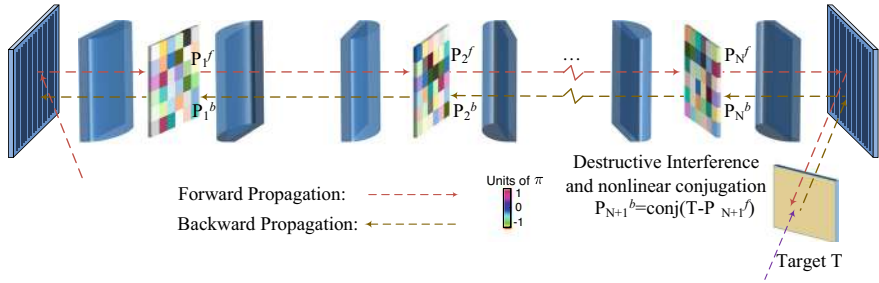


Fig. 7.24 Forward and backward propagation of the matrices $\mathbf{P}_k^f$ and $\mathbf{P}_k^b$. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

the signal is inversely injected into the STD2NN matrices $\mathbf{D}_k^{fT}$, $\mathbf{D}_k^{fT}$ and $\mathbf{M}_k^T$ are the related propagation matrices. Since the matrices are symmetric in the proposed STD2NN, one has:

$$\mathbf{D}_k^{fT} = \mathbf{D}_k^f, \mathbf{D}_k^{sT} = \mathbf{D}_k^s, \mathbf{M}_k^T = \mathbf{M}_k. \quad (7.66)$$

The conjugated error field $(\mathbf{U}_{N+1} - \mathbf{T})^*$ can be acquired through the interference between the output field $\mathbf{U}_{N+1}$ and the target field $\mathbf{T}$ followed by phase conjugation via a nonlinear process. Figure 7.24 illustrates the detailed process for the all-optical BP algorithm.

Once we calculate the gradient of the phase in each layer, the phase can be tuned as:

$$\varphi_k^{new} = \varphi_k^{old} - \mu \frac{\partial L}{\partial \varphi_k^{old}}, \quad (7.67)$$

where μ is the learning rate.

At first, the SLM's pixels can be configured with arbitrary phases, which converge to yield highly similar final patterns through iterations. For simplicity, we have set the initial phase to zero.

Although our gradient formula was derived for a single target, it can easily accommodate multiple targets by sequentially training the STD2NN on different targets.

It is important to note that, as we assume a plane wave in the x -dimension, optical computing for phase gradients only considers two dimensions. This means it requires the same computational resources as a 2D D2NN.

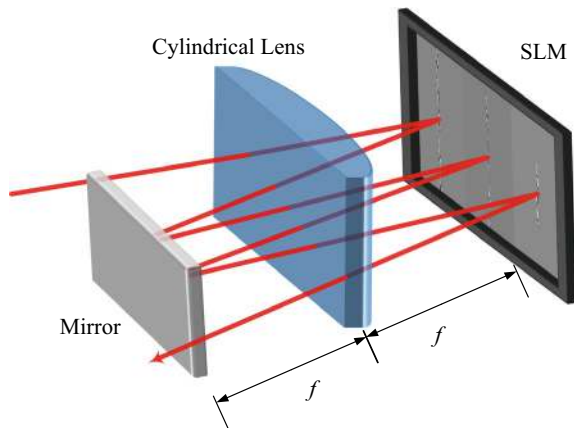
7.5.3 Results and Discussions

We have designed an STD2NN adopting the aforementioned structure. The input signal operates at a wavelength of 1550 nm. As soon as the signal enters the system, it encounters a grating with an incident angle of 50° and a diffraction angle of 9.44° . This grating features a diffraction pattern of 600 lines per mm. Following reflection, the signal passes through a cascaded pulse shaping system consisting of seven layers. Each layer comprises two lenses, both with the focal lengths of 0.005 m, and a phase-only SLM with 300×800 pixels. Each pixel measures $8 \mu\text{m} \times 8 \mu\text{m}$. Given that phase-only SLMs typically operate in reflection mode, it is possible to reuse the lens on the reflective path, and therefore, one requires only one lens per layer in experimental setups. Moreover, multiple layers can share a single SLM, as is illustrated in Fig. 7.25, where multiple $4f$ systems utilize one SLM and one cylindrical lens. While the potential experimental setup may pose calibration challenges, its feasibility has been demonstrated in [92], where a SLM and a mirror were used in a reflective configuration to generate an arbitrary spatial beam.

The initial Gaussian beam waist of the input light is $50 \mu\text{m} \times 50 \mu\text{m}$ in the x and y dimensions, matching the output beam shape of a multimode fiber. In the temporal dimension, the beam pulse also adopts a Gaussian shape. The phase optimization for the STD2NN follows Eq. (7.61). The maximum pulse duration that the STD2NN can modulate relies on the frequency frame's resolution, which is influenced by the parameter γ and pixel length. Using the parameters in this section, we calculate γ to be $3.1447\text{e-}9$. The frequency resolution per pixel is approximately 2 THz. Therefore, the maximum pulse length should be limited to 3 ps. Pulses exceeding this length are unsuitable for modulation as the STD2NN cannot resolve finer frequency units. We employ the plane wave assumption to compute both the forward error matrix $\mathbf{P}_k^f$ and the backward error matrix $\mathbf{P}_k^b$.

To test the STD2NN's capabilities, we configure it to create the word "OPTICA" in the spatiotemporal domain, as is shown in Fig. 7.23. For the OPTICA pattern,

Fig. 7.25 The potential experimental setup for the cascaded $4f$ systems. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]



we discover that a pulse duration below 0.05 ps is necessary to incorporate more high-frequency components, and therefore we assume the temporal beam waist to be 0.038 ps. It is important to note that all beam amplitudes depicted in the figures have been normalized, ensuring a maximum amplitude of one. While the x -dimension behaves as a plane wave, as is seen in Fig. 7.26a1, the Fourier plane of the 4f system applies phase modulation to the hybrid frequency-spatial plane, specifically the ω - y plane. Using “OPTICA” as the training target and optimizing the phase with 200 iterations, we achieved the outputs in Fig. 7.26b1–b2, demonstrating the generation of the word “OPTICA”.

In the possible realizations, the input beam has a waist of 50 μm in the x -dimension. By positioning it on the focal plane of an x -dimensional cylindrical lens with a focal length of 0.2 m, the beam is expanded. Two y -dimensional cylindrical lenses form a 4f system, preserving the y -dimensional beam shape. The x -dimensional beam waist is increased to approximately 2 mm, as is illustrated in Fig. 7.26a2, satisfying the plane wave assumption. Figure 7.26c1–c2 displays the spatiotemporal amplitude and phase on the y - t plane at $x = 0$ mm.

A detailed analysis is conducted to investigate the spatiotemporal beam characteristics away from the center at $x = 0$ mm, specifically exploring the validity of the plane wave assumption. This analysis revealed that the plane wave assumption remains valid for the non-center beam portion. Given that the input beam follows a Gaussian distribution in the x -dimension rather than a plane wave distribution, we

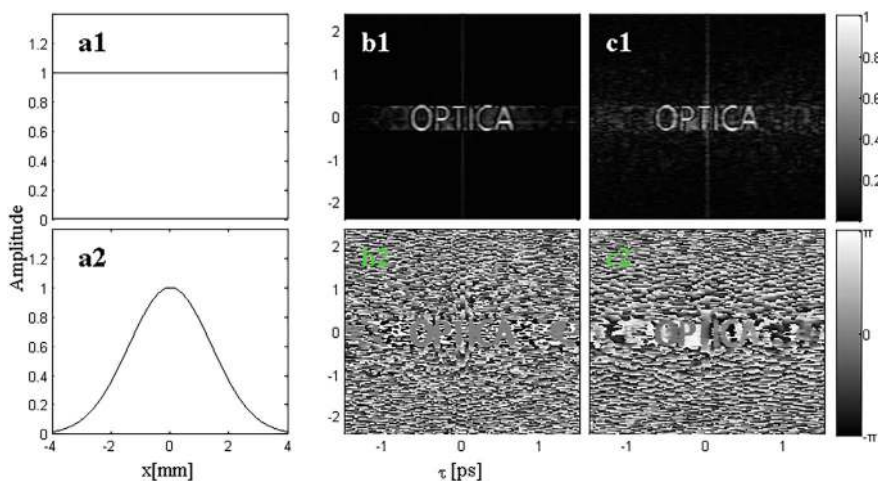


Fig. 7.26 The spatiotemporal word “OPTICA” generated by the STD2NN. **a1** The x -dimensional distribution function as 1 (the pure plane wave). **b1** The corresponding output amplitude of the STD2NN. **b2** The corresponding output phase of the STD2NN. **a2** The x -dimensional distribution function as the Gaussian function (the approximated plane wave). **c1** The corresponding output amplitude of the STD2NN. **c2** The corresponding output phase of the STD2NN. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

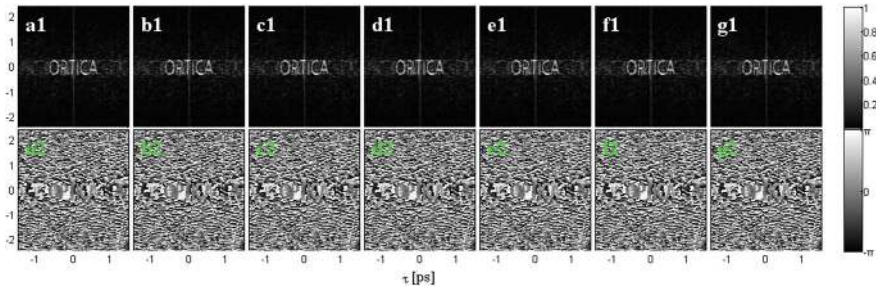


Fig. 7.27 The spatiotemporal word “OPTICA” generated by the STD2NN sliced at different x positions. **a1–g1** The spatiotemporal beam amplitudes located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. **a2–g2** The spatiotemporal beam phases located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

segmented the beam at various x -dimensional positions: $x = -3, -2, -1, 0, 1, 2$, and 3 mm. The corresponding amplitudes and phases of these spatiotemporal beams are presented in Fig. 7.27. By examining Fig. 7.27, it is evident that both the amplitude and phase remain largely unchanged across these different x -dimensional positions.

The phase distributions across the layers of the STD2NN that shape the spatiotemporal word “OPTICA” are exhibited in Fig. 7.28a–g. These distributions indicate that the phase in the center of the spatiotemporal modulation plane is crucial for shaping spatiotemporal beams. Consequently, it is feasible to reuse the phase modulation plane by positioning a mirror and an SLM together, allowing the beam to reflect between them. By inserting a cylindrical lens in the center, multiple $4f$ systems can be formed. This approach is viable because wave power is concentrated at lower frequencies, which means that only a narrow range of the SLM frequency axis is significant. The pulse duration determines the frequency range of the wave. If additional degrees of freedom are required on each phase plane, implementing a shorter pulse is recommended. Figure 7.28a1–g1 and a2–g2 depict the spatiotemporal beam amplitude and phase at the output of each layer, illustrating the transformation of an input Gaussian beam into the spatiotemporal word “OPTICA”. It’s worth noting that generating a single spatiotemporal image may require fewer diffractive layers and $4f$ systems. However, the production of multiple STOV wave-packets necessitates numerous layers, as will be demonstrated later. For consistency, we utilized the same device for the both tasks.

After showing the creation of the word “OPTICA”, we shifted our focus to a STOV wave-packet multiplexer created using the same STD2NN. The input signal wavelength and the spatial beam distribution keep unaltered, with a temporal waist of 0.19 ps. As is shown in Fig. 7.29a1, a2, two such beams are positioned 0.4 mm apart in the y -dimension. Once spatiotemporally processed by the STD2NN, these beams are transformed into two STOV wave-packets, carrying topological charges of 1 and 2 , respectively.

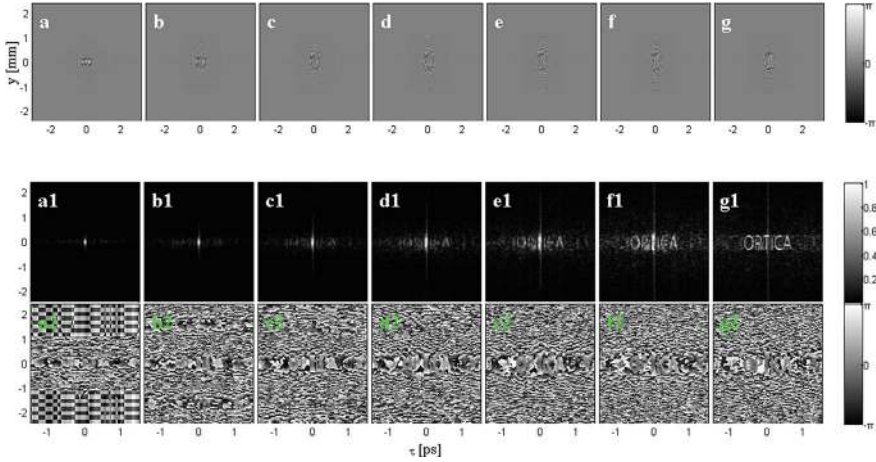


Fig. 7.28 The spatiotemporal word “OPTICA” generated by the STD2NN. **a–g** The phase distributions of the neuron layers 1–7. **a1–g1** The spatiotemporal beam amplitudes at the output of the layers 1–7. **a2–g2** The spatiotemporal beam phases at the output of the layers 1–7. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

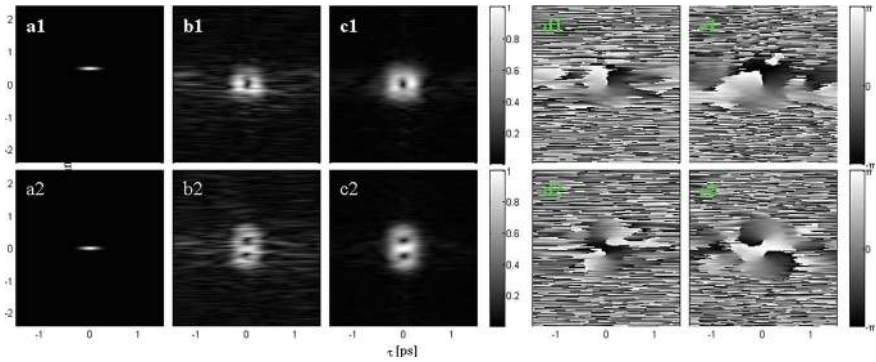


Fig. 7.29 The STOV wave-packets generated by the STD2NN. **a1** The initial spatiotemporal beam to generate STOV with $l = 1$. **b1** The STOV wave-packet amplitude with the topological charge of 1. **c1** The STOV wave-packet amplitude with the topological charge of 1 under the plane wave assumption. **d1** The STOV wave-packet phase with the topological charge of 1. **e1** The STOV wave-packet phase with the topological charge of 1 under the plane wave assumption. **a2** The initial spatiotemporal beam to generate STOV with $l = 2$. **b2** The STOV wave-packet amplitude with the topological charge of 2 under the plane wave assumption. **c2** The STOV wave-packet amplitude with the topological charge of 2 under the plane wave assumption. **d2** The STOV wave-packet phase with the topological charge of 2. **e2** The STOV wave-packet phase with the topological charge of 2 under the plane wave assumption. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

The corresponding amplitudes and phases of the beams are exhibited in Fig. 7.29b1, b2 and d1, d2. Evidently, from the figures, both STOV wave-packets coexist within the same spatiotemporal region, indicating the achievement of STOV wave-packet multiplexing. The spiral wavefronts visible in Fig. 7.29d1, d2 confirm the occurrence of STOV following the spatiotemporal transformation.

The ideal scenario, assuming a plane wave, is illustrated in Fig. 7.29c1, c2 and e1, e2. When compared to this ideal case, the actual Gaussian-shaped beam in the x -dimension does exhibit some performance degradation. However, the multiplexing effect remains clearly observable.

The x -dimensional sliced outputs from the STD2NN for STOV wave-packets are presented in Fig. 7.30. These outputs demonstrate that both the amplitude and phase remain largely unchanged across different x -dimensional positions. Figures 7.29 and 7.30 collectively suggest that STOV multiplexing can indeed be achieved using our proposed device.

The phase distributions across the layers of the STD2NN to realize STOV wave-packet multiplexing are illustrated in Fig. 7.29a–g. Once again, the active neurons are primarily clustered in the center of the spatiotemporal modulation plane, reinforcing the idea that the STD2NN, composed of cascaded 4f systems, can be implemented

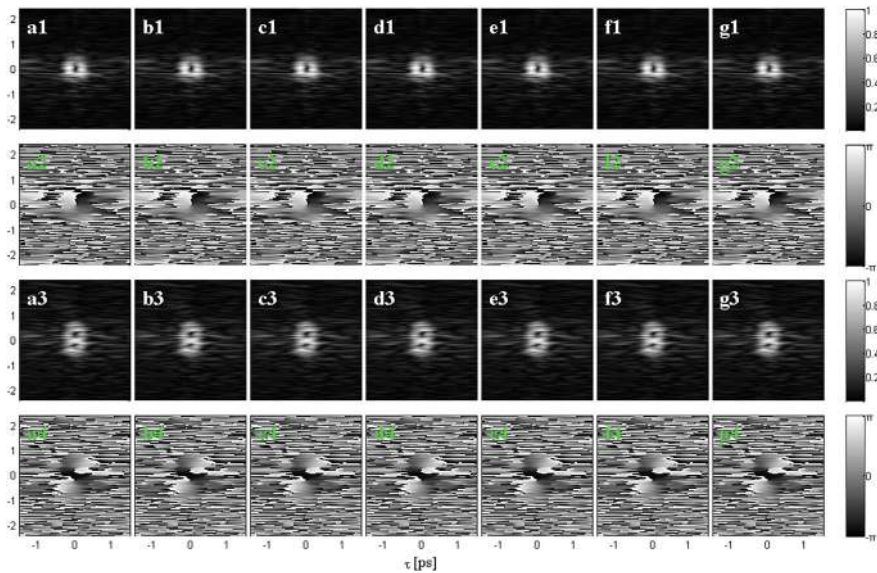


Fig. 7.30 The STOV wave-packets generated by the STD2NN sliced at different x positions. **a1–g1** The STOV wave-packet amplitudes with $l = 1$ located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. **a2–g2** The STOV wave-packet phases with $l = 1$ located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. **a3–g3** The STOV wave-packet amplitudes with $l = 2$ located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. **a4–g4** The STOV wave-packet phases with $l = 2$ located at $x = -3, -2, -1, 0, 1, 2$, and 3 mm. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

using a single SLM, as mentioned earlier. In comparison to Fig. 7.28, where the temporal beam waist is 0.76 ps, the stripe of effective neurons appears narrower in Fig. 7.31 due to a corresponding temporal beam waist of 3.8 ps.

The variation of STOV wave-packets during the conversion process is documented in Fig. 7.31a1–g1, a2–g2, a3–g3, and a4–g4. Specifically, Fig. 7.31a1–g1 and a2–g2 show the amplitudes and phases of the beams emerging from layers 1–7 during the evolution of a STOV wave-packet with a topological charge of $l = 1$. Meanwhile, Fig. 7.31a3–g3 and a4–g4 present the amplitudes and phases for the evolution of a STOV wave-packet carrying a topological charge of $l = 2$. Evidently, from these figures, we can observe the gradual transformation of the beam from its initial form without any STOV characteristics to a pure STOV wave-packet featuring a spiral wavefront with a topological charge of either $l = 1$ or $l = 2$.

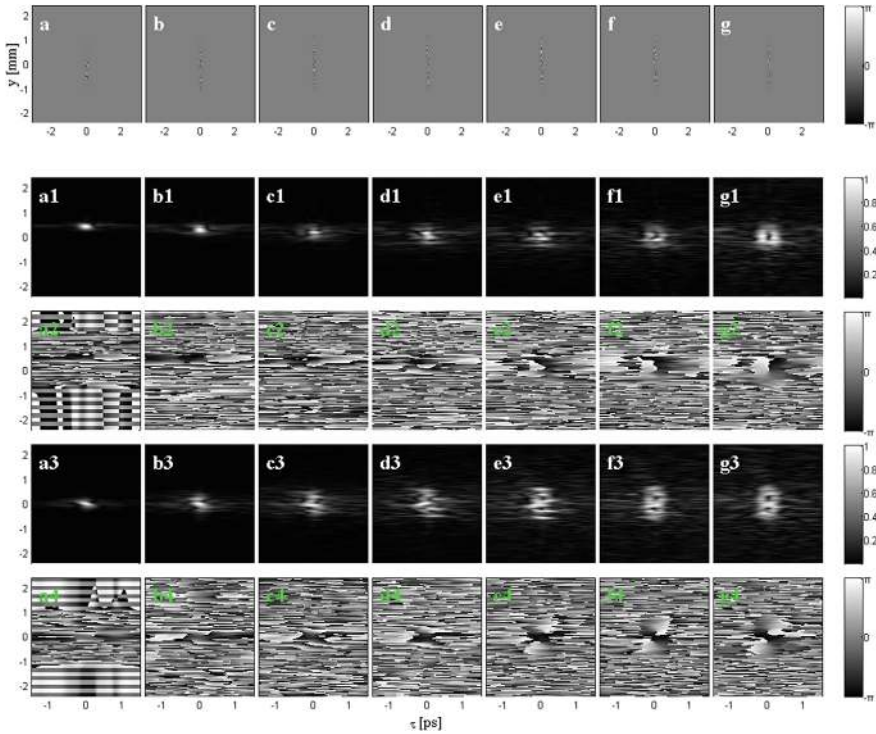


Fig. 7.31 The STOV wave-packets generated by the STD2NN at different layer outputs. **a–g** The phase distributions of the neuron layers 1–7. **a1–g1** The STOV wave-packet amplitudes with $l = 1$ at the output of the layers 1–7. **a2–g2** The STOV wave-packet phases with $l = 1$ at the output of the layers 1–7. **a3–g3** The STOV wave-packet amplitudes with $l = 2$ at the output of the layers 1–7. **a4–g4** The STOV wave-packet phases with at the output of the layers 1–7. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

Finally, we compare the proposed STD2NN with the existing spatiotemporal pulse shaping system, which consists of two gratings, a SLM, and a 4f pulse shaper [65, 83, 84]. Both the STD2NN and the traditional spatiotemporal beam shaping system aim to achieve STOV wave-packet multiplexing, which means they aim to generate STOV wave-packets with varying topological charges by injecting beams at different y positions. With the same optimization objective, the STD2NN has already exhibited its ability in Figs. 7.30 and 7.31 to accomplish STOV wave-packet multiplexing, achieving a relative error compared to the target (calculated as the integral of the loss function divided by the integral of the spatiotemporal power of the target) of approximately 35%. Meanwhile, the traditional spatiotemporal beam shaping system fails to match this performance, as is illustrated in Fig. 7.32a.

When targeting STOV wave-packets with topological charges of $l = 1$ and $l = 2$, the amplitudes and phases of the spatiotemporal beams shaped by the traditional system are presented in Fig. 7.32b1–b2 and c1–c2. Notably, while the amplitudes deviate significantly from the results observed in Figs. 7.30 and 7.31, the phases in Fig. 7.32b2 and c2 lack a spiral pattern and do not carry any topological charges. Therefore, we can infer that the traditional spatiotemporal beam shaping system is inadequate for STOV wave-packet multiplexing, whereas the proposed STD2NN proves capable of achieving this task. (The Matlab code for generating the phase of the STD2NN for STOV multiplexing is provided in Code File 1 (Ref. [93]), and the

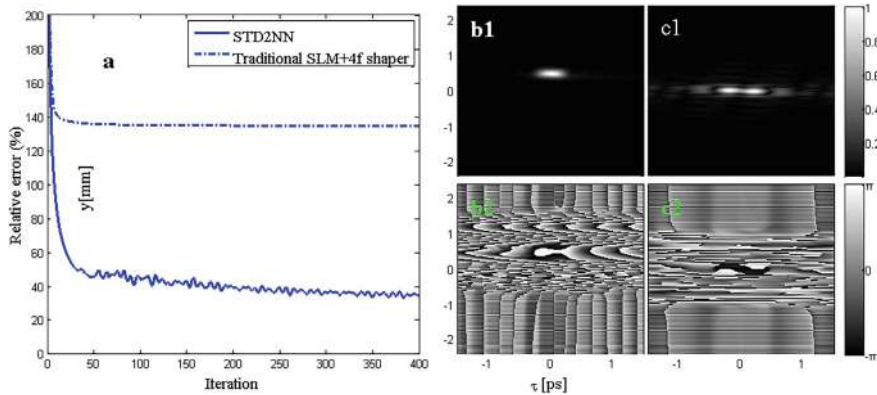


Fig. 7.32 Comparison of the results by the proposed STD2NN and the traditional spatiotemporal beam shaping system with two gratings, a SLM and a 4f pulse shaper. **a** The relative error between the output and the target after 400 neuron phase optimizations iteration by the two methods. **b1** The amplitude of the spatiotemporal beam shaped by the traditional spatiotemporal beam shaping system with the target STOV topological charge of $l = 1$. **b2** The phase of the spatiotemporal beam shaped by the traditional spatiotemporal beam shaping system with the target STOV topological charge of $l = 1$. **c1** The amplitude of the spatiotemporal beam shaped by the traditional spatiotemporal beam shaping system with the target STOV topological charge of $l = 2$. **c2** The phase of the spatiotemporal beam shaped by the traditional spatiotemporal beam shaping system with the target STOV topological charge of $l = 2$. This figure is reprinted/adapted with permission from [Opt. Express 32, 1864–1877 (2024)] © The Optical Society of America [80]

code for simulating the propagation of the 4D beam is included in Code File 2 (Ref. [93])).

7.5.4 Summary

In summary, we have introduced a novel framework for spatiotemporal signal processing, which is called STD2NN. This framework is implemented through two gratings and multiple cascaded 4f pulse shaping systems, with SLMs integrated within. We have derived an all-optical BP algorithm where the gradient of the loss function is directly proportional to the inner product of the forward field and the backward conjugated error field. We have generated the word “OPTICA” and presented a STOV wave-packet multiplexer utilizing the STD2NN as an illustration. For the experimental realization, a femtosecond pulsed laser source is necessary. The proposed STD2NN is anticipated to have a significant impact on spatiotemporal signal processing. It could potentially enable the creation of innovative optical beams, especially space–time entangled beams, and pave the way for future applications such as information multiplexing in the space–time domain. For instance, one could multiplex STOVs to carry distinct information, thereby establishing a new multiplexing paradigm.

References

1. V. Torres-Company, J. Lancis, P. Andrés, Chapter 1—Space-time analogies in optics, in *Progress in Optics*, vol. 56, ed. by E. Wolf (Elsevier, 2011), pp. 1–80
2. Y. Shen et al., Roadmap on spatiotemporal light fields. *J. Opt.* **25** (2023)
3. T. Harter et al., Generalized Kramers-Kronig receiver for coherent terahertz communications. *Nat. Photon.* **14**(10), 601–606 (2020)
4. S.A.R. Horsley, M. Artoni, G.C. La Rocca, Spatial Kramers-Kronig relations and the reflection of waves. *Nat. Photon.* **9**(7), 436–439 (2015)
5. B.W. Plansinis, W.R. Donaldson, G.P. Agrawal, What is the temporal analog of reflection and refraction of optical beams? *Phys. Rev. Lett.* **115**(18) (2015)
6. J.H. Zhou, G.Z. Zheng, J.J. Wu, Comprehensive study on the concept of temporal optical waveguides. *Phys. Rev. A* **93**(6) (2016), Art no. 063847
7. M. Born, E. Wolf, E. Hecht, Principles of optics: electromagnetic theory of propagation, interference and diffraction of light. *Phys. Today* **53**(10), 77–78 (2000)
8. M.F. Yanik, S. Fan, Time reversal of light with linear optics and modulators. *Phys. Rev. Lett.* **93**(17) (2004)
9. S.F. Preble, Q. Xu, M. Lipson, Changing the colour of light in a silicon resonator. *Nat. Photon.* **1**(5), 293–296 (2007)
10. T. Tanabe, M. Notomi, H. Taniyama, E. Kuramochi, Dynamic release of trapped light from an ultrahigh-Q nanocavity via adiabatic frequency tuning. *Phys. Rev. Lett.* **102**(4) (2009)
11. Y. Xiao, D.N. Maywar, G.P. Agrawal, Reflection and transmission of electromagnetic waves at a temporal boundary. *Opt. Lett.* **39**(3), 574–577 (2014)
12. A. Peruzzo, A. Laing, A. Politi, T. Rudolph, J.L. O’Brien, Multimode quantum interference of photons in multiport integrated devices. *Nat. Commun.* **2**(1) (2011)

13. A.R. Gupta, K. Tsutsumi, J. Nakayama, Synthesis of Hadamard transformers by use of multimode interference optical waveguides. *Appl. Opt.* **42**(15), 2730–2738 (2003)
14. Zhou, M. Zhang, All-optical discrete sine transform and discrete cosine transform based on multimode interference couplers. *IEEE Photon. Technol. Lett.* **22**(5), 317–319 (2010)
15. I. Kang et al., All-optical OFDM transmission of 7 x 5-Gb/s data over 84-km standard single-mode fiber without dispersion compensation and time gating using a photonic-integrated optical DFT device. *Opt. Express* **19**(10), 9111–9117 (2011)
16. L.M. Zhao, D.Y. Tang, H. Zhang, X. Wu, N. Xiang, Soliton trapping in fiber lasers. *Opt. Express* **16**(13), 9528–9533 (2008)
17. J.C. Travers, J.R. Taylor, Soliton trapping of dispersive waves in tapered optical fibers. *Opt. Lett.* **34**(2), 115–117 (2009)
18. J.Y.K. Aikawa, S. Saitoh, M. Kudoh, K. Suzuki, Dispersion compensating fiber module, *Fujikura Tech. Rev.* **40**(16) (2011)
19. J.-L. Auguste et al., 1800 ps/(nm.km) chromatic dispersion at 1.55 μm in dual concentric core fibre. *Electron. Lett.* **36**, 1689–1691 (2000)
20. B.J. Mangan et al., Low loss (1.7 dB/km) hollow core photonic bandgap fiber, in *Optical Fiber Communication Conference* (Los Angeles, California, 2004), Optica Publishing Group, in Technical Digest (CD), p. PD24
21. A.Y. Petrov, M. Eich, Dispersion compensation with photonic crystal line-defect waveguides. *IEEE J. Sel. Areas Commun.* **23**(7), 1396–1401 (2005)
22. T. Baba, Slow light in photonic crystals. *Nat. Photon.* **2**(8), 465–473 (2008)
23. J. Marangos, Slow light in cool atoms. *Nature* **397**(6720), 559–560 (1999)
24. J.H. Zhou, C.X. Guo, R.C. Wang, Fractional Fourier transform based on a temporal gradient-index lens. *Opt. Contin.* **1**(6), 1418–1424 (2022)
25. V. Namias, The fractional order Fourier transform and its application to quantum mechanics. *IMA J. Appl. Math.* **25**(3), 241–265 (1980)
26. G. Han, S. Li, X. Xue, X. Zheng, Photonic fractional Fourier transformer for chirp radar with ghost target elimination. *Opt. Lett.* **45**(15), 4228–4231 (2020)
27. F. Wang, Y. Wang, J. Liu, Y. Wang, Optical excitation fractional Fourier transform (FrFT) based enhanced thermal-wave radar imaging (TWRI). *Opt. Express* **26**(17), 21403–21417 (2018)
28. B. Robertson et al., Application of the fractional Fourier transform to the design of LCOS based optical interconnects and fiber switches. *Appl. Opt.* **51**(12), 2212–2222 (2012)
29. A.W. Lohmann, Image rotation, Wigner rotation, and the fractional Fourier transform. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **10**, 2181–2186 (1993)
30. D. Mendlovic, H.M. Ozaktas, Fractional Fourier transforms and their optical implementation: I. *J. Opt. Soc. Am. A* **10**(9), 1875–1881 (1993)
31. X.-B. Yang, J. Hu, Anamorphic fractional Fourier transforms graded index lens designed using transformation optics. *Opt. Express* **26**(21), 27528–27544 (2018)
32. G. Cincotti, Enhanced functionalities for AWGs. *J. Lightw. Technol.* **33**(5), 998–1006 (2015)
33. S. Weimann et al., Implementation of quantum and classical discrete fractional Fourier transforms. *Nat. Commun.* **7**(1) (2016)
34. C. Cuadrado-Laborde, A.A. Carrascosa Rubio, A. Diez, J. Cruz, M.V. Andres, Photonic fractional Fourier transformer with a single dispersive device. *Opt. Express* **21**, 8558–8563 (2013)
35. C. Schnébelin, H. Guillet de Chatellus, Agile photonic fractional Fourier transformation of optical and RF signals. *Optica* **4**(8), 907–910 (2017)
36. T. Han, H. Chen, W. Li, B. Wang, P. Lu, Temporal imaging using dispersive gradient-index time lenses. *J. Lightw. Technol.* **38**(8), 2383–2391 (2020)
37. B.D.R. Tao, Y. Wang, *Fractional Fourier Transform and Its Applications* (Tsinghua University Press, 2009)
38. J.H. Zhou, Q.S. Hu, H.Q. Pu, Nonlinear Fourier transform receiver based on a time domain diffractive deep neural network. *Opt. Express* **30**(21), 38576–38586 (2022)
39. M.I. Yousefi, F.R. Kschischang, Information transmission using the nonlinear Fourier transform, Part I: mathematical tools. *IEEE Trans. Inf. Theory* **60**(7), 4312–4328 (2014)

40. S. Gaiarin, A.M. Perego, E.P. da Silva, F. Da Ros, D. Zibar, Dual-polarization nonlinear Fourier transform-based optical communication system. *Optica* **5**(3), 263–270 (2018)
41. S.K. Turitsyn et al., Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives. *Optica* **4**(3), 307–322 (2017)
42. S.A. Derevyanko, J.E. Prilepsky, S.K. Turitsyn, Capacity estimates for optical transmission based on the nonlinear Fourier transform. *Nat. Commun.* **7**(1) (2016)
43. Y. Wu et al., Robust neural network receiver for multiple-eigenvalue modulated nonlinear frequency division multiplexing system. *Opt. Express* **28**(12), 18304–18316 (2020)
44. R.T. Jones, S. Gaiarin, M.P. Yankov, D. Zibar, Time-domain neural network receiver for nonlinear frequency division multiplexed systems. *IEEE Photon. Technol. Lett.* **30**(12), 1079–1082 (2018)
45. O. Kotlyar, M. Pankratova, M. Kamalian-Kopae, A. Vasylenkova, J.E. Prilepsky, S.K. Turitsyn, Combining nonlinear Fourier transform and neural network-based processing in optical communications. *Opt. Lett.* **45**(13), 3462–3465 (2020)
46. S. Gaiarin, F.D. Ros, N.D. Renzis, R.T. Jones, D. Zibar, Experimental demonstration of nonlinear frequency division multiplexing transmission with neural network receiver. *J. Lightw. Technol.* **38**(23), 6465–6473 (2020)
47. E.V. Sedov et al., Neural networks for computing and denoising the continuous nonlinear Fourier spectrum in focusing nonlinear Schrödinger equation. *Sci. Rep.* **11**(1) (2021)
48. E.V. Sedov, I.S. Chekhovskoy, J.E. Prilepsky, Neural network for calculating direct and inverse nonlinear Fourier transform. *Quant. Electron.* **51**(12) (2021)
49. W.Q. Zhang, T.H. Chan, S.A. Vahid, Serial and parallel convolutional neural network schemes for NFDM signals. *Sci. Rep.* **12**(1), 7962 (2022)
50. E. Mjolsness, D. DeCoste, Machine learning for science: state of the art and future prospects. *Science* **293**(5537), 2051–2055 (2001)
51. X. Lin et al., All-optical machine learning using diffractive deep neural networks. *Science* **361**(6406), 1004–1008 (2018)
52. T. Yan et al., Fourier-space diffractive deep neural network. *Phys. Rev. Lett.* **123**(2) (2019)
53. J. Bueno et al., Reinforcement learning in a large-scale photonic recurrent neural network. *Optica* **5**(6), 756–760 (2018)
54. M.S.S. Rahman, J. Li, D. Mengü, Y. Rivenson, A. Ozcan, Ensemble learning of diffractive optical networks. *Light Sci. Appl.* **10**(1), 14 (2021)
55. J. Li, D. Mengü, Y. Luo, Y. Rivenson, A. Ozcan, Class-specific differential detection in diffractive optical neural networks improves inference accuracy. *Adv. Photon.* **1** (2019)
56. T. Zhou et al., Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit. *Nat. Photon.* **15**(5), 367–373 (2021)
57. T. Zhou et al., In situ optical backpropagation training of diffractive optical neural networks. *Photon. Res.* **8**(6), 940 (2020)
58. Y. Luo et al., Computational imaging without a computer: seeing through random diffusers at the speed of light. *eLight* **2**(1), 4 (2022)
59. C.V. Bennett, B.H. Kolner, Principles of parametric temporal imaging. I. System configurations. *IEEE J. Quant. Electron.* **36**(4), 430–437 (2000)
60. J. Ashby, V. Thiel, M. Allgaier, P. d’Ornellas, A.O.C. Davis, B.J. Smith, Temporal mode transformations by sequential time and frequency phase modulation for applications in quantum information science. *Opt. Express* **28**(25), 38376–38389 (2020)
61. S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing. *Science* **220**(4598), 671–680 (1983)
62. J.H. Zhou, Q.S. Hu, Spherical Gauss-Laguerre beam propagation in 4D space-time. *Opt. Express* **31**(23), 38334–38342 (2023)
63. A. Chong, C. Wan, J. Chen, Q. Zhan, Generation of spatiotemporal optical vortices with controllable transverse orbital angular momentum. *Nat. Photon.* **14**(6), 350–354 (2020)
64. S.W. Hancock, S. Zahedpour, A. Goffin, H.M. Milchberg, Free-space propagation of spatiotemporal optical vortices. *Optica* **6**(12), 1547–1553 (2019)

65. K.Y. Bliokh, Spatiotemporal vortex pulses: angular momenta and spin-orbit interaction. *Phys. Rev. Lett.* **126**(24) (2021)
66. C. Wan, Q. Cao, J. Chen, A. Chong, Q. Zhan, Toroidal vortices of light. *Nat. Photon.* **16**(7), 519–522 (2022)
67. K.Y. Bliokh, F. Nori, Spatiotemporal vortex beams and angular momentum. *Phys. Rev. A* **86**(3) (2012)
68. N. Jhajj, I. Larkin, E.W. Rosenthal, S. Zahedpour, J.K. Wahlstrand, H.M. Milchberg, Spatiotemporal optical vortices. *Phys. Rev. X* **6**(3) (2016)
69. S.W. Hancock, S. Zahedpour, H.M. Milchberg, Second-harmonic generation of spatiotemporal optical vortices and conservation of orbital angular momentum. *Optica* **8**(5), 594–597 (2021)
70. S.W. Hancock, S. Zahedpour, H.M. Milchberg, Mode structure and orbital angular momentum of spatiotemporal optical vortex pulses. *Phys. Rev. Lett.* **127**(19) (2021)
71. J. Huang, J. Zhang, T. Zhu, Z. Ruan, Spatiotemporal differentiators generating optical vortices with transverse orbital angular momentum and detecting sharp change of pulse envelope. *Laser Photon. Rev.* **16**(5) (2022)
72. L.L. Doskolovich, A.I. Kashapov, E.A. Bezus, D.A. Bykov, Spatiotemporal optical differentiation and vortex generation with metal-dielectric-metal multilayers. *Phys. Rev. A* **106**(3), 033523 (2022)
73. M.A. Porras, Propagation of higher-order spatiotemporal vortices. *Opt. Lett.* **48**(2), 367–370 (2023)
74. S. Orlov, A. Piskarskas, A. Stabinis, Localized optical subcycle pulses in dispersive media. *Opt. Lett.* **27**(24), 2167–2169 (2002)
75. S. Longhi, Localized subluminal envelope pulses in dispersive media. *Opt. Lett.* **29**(2), 147–149 (2004)
76. O.V. Borovkova, Y.V. Kartashov, V.E. Lobanov, V.A. Vysloukh, L. Torner, General quasi-nonspredding linear three-dimensional wave packets. *Opt. Lett.* **36**(12), 2176–2178 (2011)
77. M. Dallaire, N. McCarthy, M. Piché, Spatiotemporal bessell beams: theory and experiments. *Opt. Express* **17**(20), 18148–18164 (2009)
78. O. Matsuoka, Molecular integrals over spherical Laguerre Gaussian-type functions. *J. Chem. Phys.* **92**(7), 4364–4371 (1990)
79. M. Mounaix et al., Time reversed optical waves by arbitrary vector spatiotemporal field generation. *Nat. Commun.* **11**(1), 5813 (2020)
80. J.H. Zhou, H.Q. Pu, J.X. Yan, Spatiotemporal diffractive deep neural networks. *Opt. Express* **32**(2), 1864–1877 (2024)
81. M. Yessenov et al., Space-time wave packets localized in all dimensions. *Nat. Commun.* **13**(1), 4573 (2022)
82. M. Yessenov, L.A. Hall, K.L. Schepler, A.F. Abouraddy, Space-time wave packets. *Adv. Opt. Photon.* **14**(3), 455–570 (2022)
83. S. Huang, P. Wang, X. Shen, J. Liu, Properties of the generation and propagation of spatiotemporal optical vortices. *Opt. Express* **29**(17), 26995–27003 (2021)
84. H. Wang, C. Guo, W. Jin, A.Y. Song, S. Fan, Engineering arbitrarily oriented spatiotemporal optical vortices using transmission nodal lines. *Optica* **8**(7), 966–971 (2021)
85. M.M. Wefers, K.A. Nelson, Space-time profiles of shaped ultrafast optical waveforms. *IEEE J. Quant. Electron.* **32**(1), 161–172 (1996)
86. H.E. Kondakci, A.F. Abouraddy, Diffraction-free space–time light sheets. *Nat. Photon.* **11**(11), 733–740 (2017)
87. L. Chen et al., Synthesizing ultrafast optical pulses with arbitrary spatiotemporal control. *Sci. Adv.* **8**(43) (2022)
88. J. Zhou, Q. Hu, H. Pu, Nonlinear Fourier transform receiver based on a time domain diffractive deep neural network. *Opt. Express* **30**(21), 38576–38586 (2022)
89. M.S.S. Rahman, A. Ozcan, Time-lapse image classification using a diffractive neural network. *Adv. Intell. Syst.* **5**(5), 2200387 (2023)
90. T. Zhou, W. Wu, J. Zhang, S. Yu, L. Fang, Ultrafast dynamic machine vision with spatiotemporal photonic computing. *Sci. Adv.* **9**(23) (2023)

91. R.J. Potton, Reciprocity in optics. *Rep. Progr. Phys.* **67**(5), 717 (2004)
92. N.K. Fontaine, R. Ryf, H. Chen, D.T. Neilson, K. Kim, J. Carpenter, Laguerre-Gaussian mode sorter. *Nat. Commun.* **10**(1) (2019)
93. J. Zhou, H. Pu, J. Yan, Spatiotemporal diffractive deep neural networks. *Opt. Express* **32**(2), 1864–1877 (2024)

Index

C

Coupled mode theory, 1, 5, 45, 48–50, 57–59, 61, 86, 94, 97–99, 102, 108, 109, 137, 138, 157, 158, 174, 177, 188, 189, 197, 201, 213
Coupled theory, 188

F

Free space, 1, 2, 13, 15, 20–22, 26, 28, 30, 32–34, 46, 47, 85, 98, 110, 124, 137, 139, 144, 197–200, 213–216, 237, 238, 240, 246–249, 254, 256–258, 260, 281, 288, 308, 338–341
Free space propagation, 137

M

Mode division multiplexing, 138, 159, 237, 261
Mode generation, 103, 237, 256
Mode generator, 237
Mode multiplexers, 13, 97, 237, 248, 261, 280
Multicore fibers, 174, 175, 188, 261
Multi-mode in space and time, 297, 307, 327, 337
Multi-mode interference, 1, 9, 11, 13, 20, 21, 32, 97, 121, 237–239, 245, 246, 248, 262, 298, 299, 303, 306
Multi-mode waveguide, 77, 137, 237

O

Optical waveguide, 21, 32, 45, 46, 48, 49, 57, 65, 137–139, 151, 157, 198, 199, 201, 203, 297, 298, 300

P

Physics, 1, 9, 62, 78, 297

Q

Quantum interference, 48, 49, 57
Quantum optics, 45, 48, 49, 57, 58, 303

R

Random coupling, 137, 138, 158, 159, 175–177, 181, 183, 184, 186–188, 194–197

S

Spatial-temporal diffractive deep neural network, 297
Spatiotemporal signals, 337–339, 341, 351, 352
Stochastic differential equation, 137, 139, 153, 158, 159, 163, 188, 191
Stochastic equation, 137, 139, 142, 153, 158, 159, 163, 175, 188, 189, 191
Super-modes, 77–80, 83, 85, 86, 94–98, 100–105, 107–109, 112, 115, 117, 119, 121, 174–176, 178, 181–186, 189, 194, 237–239, 243, 245–247

T

Temporal waveguide, 297–307