Digital Signal Processing Algorithms in Single-Carrier Optical Coherent Communications

Thèse

Wing-Chau Ng

Doctorat en génie éléctrique
Philosophiæ doctor (Ph.D.)

Québec, Canada

Résumé

Des systèmes de détection cohérente avec traitement numérique du signal (DSP) sont présentement déployés pour la transmission optique de longue portée. La modulation par déplacement de phase en quadrature à deux polarisations (DP-QPSK) est une forme de modulation appropriée pour la transmission optique sur de longues distances (1000 km ou plus). Une autre forme de modulation, le DP-16-QAM (modulation d'amplitude en quadrature) a été récemment utilisée pour les communications métropolitaines (entre 100 et 1000 km). L'extension de la distance maximum de transmission du DP-16-QAM est un domaine de recherche actif. Déterminer si l'utilisation de la détection cohérente pour les transmissions à courtes distances (moins de 100 km) en justifieraient les coûts demeure cependant une question ouverte. Dans cette thèse, nous nous intéresserons principalement au recouvrement de phase et au démultiplexage en polarisation dans les récepteurs numériques cohérents pour les applications à courte distance.

La réalisation de systèmes optiques gigabauds cohérents en temps-réel utilisant des formats de modulation à monoporteuse plus complexes, comme le 64-QAM, dépend fortement du recouvrement de phase. Pour le traitement numérique hors-ligne, la récupération de phase utilisant les résultats de décisions (decision-directed phase recovery (DD-PR)) permet d'obtenir, au débit des symboles, les meilleures performances, et ce avec un effort computationnel moindre que celui des meilleurs algorithmes connus. L'implémentation en temps-réel de systèmes gigabauds requiert un haut degré de parallélisation qui dégrade de manière significative les performances de cet algorithme. La parallélisation matérielle et le délais de pipelinage sur la boucle de rétroaction imposent des contraintes strictes sur la largeur spectrale du laser, ainsi que sur le niveau de bruit spectral des sources laser. C'est pourquoi on retrouve peu de démonstrations de recouvrement de phase en temps-réel pour les modulations 64-QAM ou plus complexes. Nous avons analysé expérimentalement l'impact des lasers avec filtres optiques sur le recouvrement de phase realisé en pipeline sur un système cohérent à monoporteuse 64-QAM à 5 Gbaud. Pour les niveaux de parallélisation plus grands que 24, le laser avec filtres optiques a permis une amélioration de 2 dB du ratio signal-à-bruit optique, en comparaison avec le même laser sans filtre optique.

La parallélisation du recouvrement de phase entraîne non seulement une plus grande sensibilité

au bruit de phase du laser, mais aussi une plus grande sensibilité aux fréquences résiduelles induites par la présence de tonalités sinusoïdales dans la source. La modulation de fréquences sinusoïdales peut être intentionnelle, pour des raisons de contrôle, ou accidentelles, dues à l'électronique ou aux fluctuations environnementales. Nous avons étudié expérimentalement l'impact du bruit sinusoïdal de phase du laser sur le système parallèle de recouvrement de phase dans un système 64-QAM à 5 Gbauds, en tenant compte des effets de la compensation du décalage de fréquence et de l'égalisation.

De nos jours, les filtres MIMO (multi-input multi-output) à réponse finie (FIR) sont couramment utilisés pour le démultiplexage en polarisation dans les systèmes cohérents. Cependant, ces filtres souffrent de divers problèmes durant l'acquisition, tels la singularité (les mêmes données apparaissent dans les deux canaux de polarisation) et la longue durée de convergence de certaines combinaisons d'états de polarisation (SOP). Pour réduire la consommation d'énergie exigée dans les systèmes cohérents pour les applications à courtes distances où le délais de groupe différentiel n'est pas important, nous proposons une architecture DSP originale. Notre approche introduit une pré-rotation de la polarisation, avant le MIMO, basée sur une estimation grossière de l'état de polarisation qui n'utilise qu'un seul paramètre Stokes (s_1). Cette méthode élimine les problèmes de singularité et de convergence du MIMO classique, tout en réduisant le nombre de filtres MIMO croisés, responsables de l'annulation de la diaphonie de polarisation. Nous présentons expérimentalement un compromis entre la réduction de matériel et la dégradation des performances en présence de dispersion chromatique résiduelle, afin de permettre la réalisation d'applications à courtes distances.

Finalement, nous améliorons notre méthode d'estimation à l'aveugle par un filtre Kalman étendu (EKF) à temps discret de faible complexité, afin de réduire la consommation de mémoire et les calculs redondants apparus dans la méthode précédante. Nous démontrons expérimentalement que la pré-rotation de polarisation basée sur le EKF operé au taux ASIC (Application-Specific Integrated Circuits) permet de récupérer la puissance de fréquence d'horloge du signal multiplexé en polarisation ainsi que d'améliorer la performance du taux d'erreur sur les bits (BER) en utilisant un MIMO de complexité réduite.

Abstract

Coherent detection with digital signal processing (DSP) is currently being deployed in long-haul optical communications. Dual-polarization (DP) quadrature phase shift keying (QPSK) is a modulation format suitable for long-haul transmission (1000 km or above). Another modulation, DP-16-QAM (quadrature amplitude modulation) has been deployed recently in metro regions (between 100 and 1000 km). Extending the reach of DP-16QAM is an active research area. For short-reach transmission (shorter than 100 km), there is still an open question as to when the technology will be mature enough to meet cost pressures for this distance. In this dissertation, we address mainly on phase recovery and polarization demultiplexing in digital coherent receivers for short-reach applications.

Implementation of real-time Gbaud (Gsymbol per second) optical coherent systems for single-carrier higher-level modulation formats such as 64-QAM depends heavily on phase tracking. For offline DSP, decision-directed phase recovery is performed at the symbol rate with the best performance and the least computational effort compared to best-known algorithms. Real-time implementations at Gbaud requires significant parallelizing that greatly degrades performance of this algorithm. Hardware parallelization and pipelining delay on the feedback path impose stringent requirements on the laser linewidth, or the frequency noise spectral level of laser sources. This leads to the paucity of experiments demonstrating real-time phase tracking for 64- or higher QAM. We experimentally investigated the impact of optically-filtered lasers on parallel and pipelined phase tracking in a single-carrier 5 Gbaud 64-QAM back-to-back coherent system. For parallelization levels higher than 24, the optically-filtered laser shows more than 2 dB improvement in optical signal-to-noise ratio penalty compared to that of the same laser without optical filtering.

In addition to laser phase noise, parallelized phase recovery also creates greater sensitivity to residual frequency offset induced by the presence of sinusoidal tones in the source. Sinusoidal frequency modulation may be intentional for control purposes, or incidental due to electronics and environmental fluctuations. We experimentally investigated the impact of sinusoidal laser phase noise on parallel decision-directed phase recovery in a 5 Gb 64-QAM system, including the effects of frequency offset compensation and equalization.

MIMO (multi-input multi-output) FIR (finite-impulse response) filters are conventionally used

for polarization demultiplexing in coherent communication systems. However, MIMO FIRs suffer from acquisition problems such as singularity and long convergence for a certain polarization rotations. To reduce the chip power consumption required in short-reach coherent systems where differential group delay is not prominent, we proposed a novel parallelizable DSP architecture. Our approach introduces a polarization pre-rotation before MIMO, based on a very-coarse blind SOP (state of polarization) estimation using only a single Stokes parameter (s_1). This method eliminates the convergence and singularity problems of conventional MIMO, and reduces the number of MIMO cross taps responsible for cancelling the polarization crosstalk. We experimentally presented a tradeoff between hardware reduction and performance degradation in the presence of residual chromatic dispersion for short-reach applications.

Finally, we extended the previous blind SOP estimation method by using a low-complexity discrete-time extended Kalman filter in order to reduce the memory depth and redundant computations of the previous design. We experimentally verified that our extended Kalman filter-based polarization prerotation at ASIC rates enhances the clock tone of polarization-multiplexed signals as well as the bit-error rate performance of using reduced-complexity MIMO for polarization demultiplexing.

Table des matières

Liste des tableaux

Liste des figures

To my mother, Kam-Ying Lee.

Acknowledgement

First and foremost, I would like to express my gratitude to my advisor, Prof. Leslie A. Rusch for her support, resources, advice, experience and guidance in the completion of this work.

I would also show my thank to Dr. An T. Nguyen, of Université Laval, who is responsible for experiments, equipment, signal generation, transmitter signal equalization, fiber optics, advices and feedback on my digital signal processings.

Many thanks and much appreciation are owed to Dr. Chul-Soo Park, of Université Laval, who is a very professional laboratory manager and research scientist in our optical communication laboratory, responsible for experiment, equipment, signal generation, transmitter and receiver hardware, advices and guidance on the equipment handling and maintenance.

I also thank our laboratory technician Mr. Philippe Chrétien for his great help in providing speedy support in laboratory tools, logistics, safety guidance, equipment handling.

I would like to thank Dr. Qunbi Zhuge of McGill University for technical discussion in our countless email correspondances. His explaination on his paper about parallelization and superscalar structure for phase tracking contributed a lot in my first project in Chapter 3.

I also express my gratitude to the following persons, without priority, for their fruitful technical discussion on timing phase recovery in coherent communications : Yuliang Gao of Hong Kong Polytechnic University, Meng Qiu, Xian Xu and Wei Wang of McGill University, which contributes a lot of my understanding and leads to my work in Chapter 6 and in Appendix G.

I would also thank Dr. Peter Winzer of Bell Labs, New Jersey, for being responsible and active for replying emails even though I am only an unimportant figure in the research field ; Dr. Xi Chen of University of Melbourne for her email correspondence regarding the experimental setup of laser linewidth characterization ; Dr. Kai Shi of Dublin City University for his email correspondence regarding the experimental setup of laser linewidth characterization ; Dr. Robert Maher of University College of London for his email correspondence regarding his paper on dynamic linewidth measurement ; Dr. F. Vacondio of Bell Labs, Paris, for notifying me the paper from Kikuchi [26] in an email correspondence in 2012.

Special thanks are also given to the colleague, Mr. Charles Brunet, for his time to translate my thesis abstract from English to French, and to the following past colleagues : Dr. Amirhossein Ghazisaeidi, for his political correctness and generous enlightenment in my system work ; Dr. Qing Xu, for his countless extra hours helping me finish my experiment during my visit in May 2010 ; Dr. Mehrdad Mirshafiei, for his unlimited information about the living in Québec City and for his friendly support.

Upon the completion of the final version of this dissertation, I would thank Mr. Jiachuan Lin and Mr. Zhihui Cao, of Université Laval for their prompt responses and their help in coordinating computers for me to compute the results required.

Last but not least, I express my gratitude to my mother and my sister for their patience, spiritual and financial support to let me study further after the completion of my master of applied science degree at University of Toronto ; to Prof. Chester Shu, Prof. Hong-Ki Tsang and Prof. Kong-Pan Poon, of the Chinese University of Hong Kong, for their support and opinions on my graduate studies, and to my secondary-school teacher, Mr. Wai-Shing Chan, for his enlightenment and rational advice for every decision I made for my life.

Avant-propos

Publications related to this thesis

1. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch, "Overcoming Phase Sensitivity in Real-time Parallel DSP for Optical Coherent Communications : Optically Filtered Lasers," *IEEE J. Lightw. Technol.*, **Vol**. 32, **No**. 3, pp. 411 - 420 Feb. 2014. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6678710

   The problem (focusing on parallelized phase tracking) was defined by L. A. Rusch. Dr. Simon Ayotte of TeraXion collaborated with us and provided access to narrow linewidth TeraXion sources for use in our experimental demonstrations. The aim of this paper is to experimentally investigate the regimes where the TeraXion narrow-linewidth lasers outperform conventional lasers.

   I was responsible for digital signal processing (DSP) and data analysis. A. T. Nguyen and C.-S. Park were responsible for the experimental setup. The first experiment of back-to-back 16-QAM was first set up by Chul-Soo Park in August/September 2012 for bit-error rate measurement for M-QAM for a conference paper [39]. In March 2013, An T. Nguyen performed an experiment of predistortion of 64-QAM transmitted data. I performed the measurement of the bit error rate (BER) versus optical signal-to-noise ratio (OSNR). Finally, I wrote the entire paper, with many helpful comments/suggestions from all coauthors. Please refer to Chapter 3.

2. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch, "Impact of Sinusoidal Tones on Parallel Decision-Directed Phase Recovery for 64-QAM," *IEEE Photonics Letters Technology*, **Vol**. 26, **No**. 5, pp. 486 - 489, Mar. 2014. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6701345

   I proposed this research topic, finding out the parameters of sine tones (for control purposes) of laser sources that do not result in failure of parallel phase tracking algorithm with feedback delay.

   I was responsible for all DSP and data analysis. An T. Nguyen and Chul-Soo Park were responsible for experimental setup. An T. Nguyen performed the 64-QAM predistortion. I performed the measurement of the bit error rates. All coauthors provided valuable comments/suggestions on the text. Please refer to Chapter 4.

3. **W. C. Ng**, A. T. Nguyen, C. S. Park, and L. A. Rusch, "Reduction of MIMO-FIR Taps via SOP-Estimation in Stokes Space for 100 Gbps Short Reach Applications," *European Conference and Exhibition on Optical Communication*, P3.3, September, 2014. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6963901

   I proposed this research topic, a novel DSP architecture to accelerate algorithm convergence and restore the clock tone to avoid system failure in a single-carrier coherent receiver. The architecture reduces the power consumption (hardware) of the single-carrier coherent receiver chips for short-reach application.

   I was responsible for algorithm design, digital signal processing (DSP) and data analysis. An T. Nguyen and Chul-Soo Park were responsible for experimental setup. The experiment of 40-km single-mode fiber (SMF) 32 Gbaud DP-QPSK was set up by An T. Nguyen, with the assistance of Chul-Soo Park.Finally, I wrote the entire paper, with many helpful comments and suggestions from all coauthors. Please refer to Chapter 6.

4. **W. C. Ng**, A. T. Nguyen, C. S. Park, and L. A. Rusch, "Enhancing Clock Tone via Polarization Pre-rotation : A Low-complexity, Extended Kalman Filter-based Approach,"*Optical Fiber Communication Conference*, paper Th2A.19, March 2015.

   I proposed this research topic, a low-complexity algorithm using the extended Kalman filter to further reduce the hardware consumption required in the previous work.

   I was responsible for algorithm design, DSP and data analysis. An T. Nguyen and Chul-Soo Park were responsible for the experimental setup. The experiment of 40-km single-mode fiber (SMF) 32Gbaud DP-QPSK was set up by An T. Nguyen in March/April 2014, with the assistance of Chul-Soo Park. Finally, I wrote the entire paper, with many helpful comments and suggestions from all coauthors. Please refer to Chapter 7.

Abbreviations

ADC	Analog-to-Digital Conversion/Converter
ASIC	Application-Specific Integrated Circuits
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BPS	Blind Phase Search
CD	Chromatic Dispersion
CDC	Chromatic Dispersion Compensation
CM	Complex Multiplication/Multipliers
CMA	Constant Modulus Algorithm
CMOS	Complementary Metal-Oxide-Semiconductor
CPE	Carrier Phase Estimation
CT	Clock Tone
DD-LMS	Decision-Directed Least-Mean Squares
DD-MLE	Decision-Directed Maximum Likelihood Estimation
DD-PR	Decision-Directed Phase Recovery
DGD	Differential Group Delay
DP	Dual-Polarization
DSP	Digital Signal Processing
ECL	External Cavity Laser
EDFA	Erbium-Doped Fiber Amplifier
EKF	Extended Kalman Filter
FBG	Fiber Bragg Grating
FD	Frequency-Domain
FEC	Feedforward Error Correction
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FN-PSD	Frequency-Noise Power Spectral Density
FO	Frequency Offset
FOC	Frequency Offset Compensation

I-Q	Inphase-Quadrature-Phase
IM-DD	Intensity Modulation with Direct Detection
ISI	Inter-symbol Interference
ITLA	Integrated Tunable Laser Assembly
LHS	Left-Hand Side
LMS	Least-Mean Squares
MIMO	Multi-Input Multi-Output
MMSE	Minimum Mean Squared Error
MSE	Mean Squared Error
MQAM	M-ary Quadrature Amplitude Modulation
MQPSK	M-ary Quadrature Phase Shift Keying
NDA	Non-Data-Aided
OOK	On-Off Keying
OSA	Optical Spectrum Analyzer
OSNR	Optical Signal-to-Noise Ratio
PMD	Polarization Mode Dispersion
PN-PSD	Phase-Noise Power Spectral Density
PolDemux	Polarization Demultiplexing
PolMux	Polarization Multiplexing
PR	Pre-rotation
PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RHS	Right-Hand Side
RM	Real Multiplication/Multiplier
RTO	Real-Time Oscilloscope
SMF	Single-Mode Fiber
SOP	State-Of-Polarization
SPS	Samples Per Symbol
SS	Stokes-Space
TD	Time-Domain
TPE	Timing Phase Estimation
WH	Wiener-Hopf
WH-DD-EQ	Wiener-Hopf Decision-Directed EQualization/EQualizer

List of Symbols by Chapter

Chapter 2 :

f_{TX}	Center frequency of transmit laser
f_{RX}	Center frequency of receive laser
Δf	Frequency offset between transmit and receive lasers
$\theta_{\mathrm{TX}}(t)$	Phase noise of transmit laser
$\theta_{\mathrm{RX}}(t)$	Phase noise of receive laser
$\theta_{LPN}(t)$	Total laser phase noise (including both transmit and receive lasers)
$d(t)$	Transmitted data
$I(t)$	Real part of transmitted data
$Q(t)$	Imaginary part of transmitted data
$n_{ASE}(t)$	Complex amplified spontaneous emission noise
$n'_{ASE}(t)$	Rotated complex amplified spontaneous emission noise
k	Discrete time index
T_{sym}	Symbol duration
d_k	Transmitted data after sampling
r_k	Received signal after sampling
n_k	Additive white Gaussian noise after sampling
θ_k	Total laser phase noise after sampling
$\hat{d}_k$	Estimate of transmitted data
$\hat{\theta}_k$	Estimate of total phase phase noise
$p(\cdot)$	Probability density function
σ	Standard deviation of the AWGN noise n_k
P_k	M-QAM signal power
V	Phaser (used in decision-directed maximum likelihood phase estimation)
M	Order or constellation size of PSK or QAM signals
B	Number of test phase candidates

Chapter 3 and Appendix D :

φ_b	Test phase angle for blind phase search
$\hat{X}_{k,b}$	Output of the decision device for blind phase search
Y_k	The sampled received signal for blind phase search
P	Number of parallelization
d	Number of pipelining delay elements
U	Phase increment
$S_{\Delta f}$	Frequency noise power spectral density level
$S_{\Delta f,BM}$	Frequency noise power spectral density level due to Brownian-motion laser phase noise
$S_{\Delta f,AWGN}$	Frequency noise power spectral density level due to the AWGN noise from electronics or EDFA
A	Amplitude of electric field of the receiver signal in the absence of noise
N_o	Two-sided field power spectral density of the AWGN noise
T_{loop}	Response time of the feedback loop of parallel DD-MLE
R_s	Symbol rate
K	Proportionality constant

Chapter 4 :

$\Delta\hat{f}_{FO}$	Frequency offset estimate
L_{FFT}	Block size of fast Fourier transform
f_c	Laser carrier frequency
A_{pp}	Peak-to-peak frequency-modulated amplitude
f_m	Frequency-modulated frequency
L_{DS}	Length of data segment used to undergo FOC

Chapter 5 :

θ	Polarization orientation angle
ϕ	Phase delay between the optical fields on X and Y polarizations
τ	Differential group delay (DGD) due to polarization mode dispersion (PMD)
X_{in}^k	Input sequence of MIMO for X-polarization
Y_{in}^k	Input sequence of MIMO for Y-polarization
X_{out}	Output of MIMO (after equalization and polarization de-rotation) for X-polarization
Y_{out}	Output of MIMO (after equalization and polarization de-rotation) for Y-polarization
$\underline{h}_{i,j}^k$	FIR filters in MIMO, where i and j can be x or y
μ	Step-size parameter for updating adaptive FIR filters
∇J	Averaged error signal depending on the criterion used for equalization
∇J_{CMA}	Averaged error signal based on CMA
$\underline{A}$	Electric field or vector on either X or Y-polarization
e_x	Instantaneous error signal based on CMA, calculated based on the X output of MIMO
e_y	Instantaneous error signal based on CMA, calculated based on the Y output of MIMO

Chapter 6 :

$\hat{s}$	Estimated state-of-polarization
s_i	Stokes parameters, where i can be 1, 2 or 3
$\hat{\theta}$	Estimated polarization orientation angle
$\hat{\phi}$	Estimated phase delay between the optical fields on X and Y polarizations
E_x	Complex field in X-polarization
E_y	Complex field in Y-polarization

Chapter 7 :

$\underline{x}_k$	State vector
$\underline{\hat{x}}_k$	Posteriori estimate of $\underline{x}_k$
$\underline{\hat{x}}_k^-$	Priori estimate of $\underline{x}_k$
ζ_{k-1}	Process noise
u_k	External perturbation
ξ_k	Measurement noise
z_k	Measurement output
h	Measurement system
Q	Process-noise covariance variance
R	Measurement-noise covariance variance
H	Jacobian matrix of partial derivatives of function h
K_k	Kalman gain
P_{k-1}	Covariance matrix of estimation error
P_k^-	Covariance matrix of estimation error after the adjustment using the knowledge of parameter dynamics A and process noise Q
x_0	Initial state
m_0	Mean of initial state x_0
P_0	Covariance matrix of initial state x_0
J_{PR}	Jones matrix for pre-rotation
X_{in}	X-polarization input of J_{PR}
Y_{in}	Y-polarization input of J_{PR}
X_{out}	X-polarization output of J_{PR}
Y_{out}	Y-polarization output of J_{PR}

Appendix B and Appendix F :

$\Delta f(t)$	Frequency noise of a laser source
$\sigma^2_{\Delta f}$	Variance of frequency noise
$S_{\Delta f}$	Power spectral density of frequency noise
$\theta(t)$	Laser phase noise
$\theta_\tau(t)$	Phase noise increment over time interval τ
$\sigma^2_{\theta_\tau}$	Variance of $\theta_\tau(t)$
$\Delta\nu$	Total laser linewidth
θ_k	Discretized laser phase noise
f_k	Discretized laser frequency noise
$\Sigma^2_{\Delta f}$	Variance of discretized laser frequency noise
T_s	Sampling duration
BW_s	Sampling bandwidth or simulation bandwidth

Appendix C :

P_{CMOS}	Power consumption of a CMOS circuit
C_{total}	Total capacitance of a circuit
V_0	Supply voltage
f_{clk}	Clock frequency of a circuit
T_{proc}	Minimum allowed clock period of a processor
V_t	CMOS threshold voltage
C_{charge}	Capacitance to be charged or discharged in a single clock cycle
k	Process parameter dependent on the material
	and the geometry applied in the CMOS techonology
$T_{pd,s,pip}$	Propagation delay for pipelined serial processing
$C_{charge,pip}$	Charging capacitance after pipelining
$T_{pd,p,pip}$	Propagation delay for pipelined processing
	at each parallelization channel
$P_{CMOS,p,pip}$	Power consumption of a CMOS circuit with parallel and pipelined processing
P	Number of parallelization

Appendix E :

$BW_{2-sided}$	Two-sided bandwidth of the spectrum or sampling rate simulation bandwidth
N_o	Two-sided field power spectral density of the AWGN noise
A	Amplitude of electric field of the receiver signal in the absence of noise
r_k	Received signal after sampling
θ_k	Total laser phase noise after sampling
$S_{\Delta f,AWGN}$	Frequency noise power spectral density level due to the AWGN noise from electronics or EDFA

Chapitre 1

Introduction

1.1 Motivation

Complex modulation with coherent detection system offers much greater spectral efficiency than intensity modulation with direct detection (IM-DD). Using coherent detection, both amplitude and phase can be retrieved at the receiver, enabling digital signal processing (DSP) to compensate chromatic and polarization mode dispersion. DSP can achieve polarization demultiplexing, enabling dual-polarization transmission systems that double the transmission capacity. Several factors contribute to the recent revival of coherent communications, such as the advancement of analog-to-digital converters, dual-polarization integrated coherent receivers and low-linewidth lasers.

Coherent systems have been deployed in long-haul optical communications in recent years. Dual-polarization (DP) quadrature phase shift keying (QPSK) is a standard modulation format suitable for long-haul transmission (1000 km and higher) due to its robustness to additive white Gaussian noise (AWGN). Another modulation, DP-16 QAM (quadrature amplitude modulation) has been deployed recently in metro region (100 - 1000 km). Extending the reach of DP-16QAM is under investigation for metro networks. For short-reach transmission (shorter than 100 km), there is still an open question regarding the migration of mature coherent technology into this area. This thesis focuses on DSP in coherent receivers for short-reach applications. In particular, we will investigate how we can improve carrier phase recovery, frequency offset estimation, polarization demultiplexing and timing phase recovery.

1.2 Objectives and Contributions

1.2.1 Problem Statement for Chapter 3 and Chapter 4

In optical coherent communications, phase noise of laser sources is a major impairment even in the absence of an optical channel. Phase recovery at the symbol rate is required to esti-

mate and compensate the fast-varying random laser phase noise, and can only be performed using DSP. The phase recovery for QPSK and 16-QAM requires simple feedforward algorithms such as Viterbi-Viterbi algorithms or blind phase search with straightforward hardware parallelization, to achieve theoretical optimal performance. For 64-QAM, decision-directed phase recovery algorithm is preferred because of the higher accuracy required. However, feedback delay in decision-directed phase recovery algorithm is the major problem when considering hardware parallelization and pipelining in implementation. Feedback algorithms were implemented in offline DSP in numerous 64-QAM experimental demonstration, without taking into account feedback delay, leading to over-optimistic results. Moreover, the existing literature adheres to the use of laser linewidths as a figure-of-merit for quantifying laser phase noise. This is only reasonable for analytically evaluating system performance with phase recovery carried out at symbol rates, assuming laser phase noise following Brownian motion (Wiener process). When the laser phase noise is non-Wiener or the laser phase is modulated by sinusoidual waves, called sine tones (which are caused by mechanical vibrations or by electronics for control purposes), this figure-of-merit can no longer be justified.

1.2.2 Contributions for Chapter 3 and Chapter 4

In my first project (Chapter 3), we showed that the conventional method of using laser linewidth for phase-noise quantification is inappropriate when laser phase noise is non-Wiener, or when feedback algorithms are used for phase recovery. Instead, we applied the frequency noise power spectral density level as a tool to quantify the impact of laser phase noise for real-time phase tracking with feedback delay. Our analysis went beyond offline experimental demonstrations assuming the absence of feedback delay. We experimentally demonstrated that TeraXion's novel narrow-linewidth laser sources improve system performance, compared to that of conventional laser sources, when taking into account the feedback delay in phase recovery. We found that the frequency noise of laser sources between 10 MHz and 100 MHz affects the real-time decision-directed phase recovery as a result of reduction of tracking bandwidth due to hardware parallelization and pipelining. This result should hold for higher symbol rates as well, as a higher symbol rate requires higher parallelization in the current CMOS technology. For parallelization level higher than 24, TeraXion's narrow-linewidth lasers show more than 2 dB improvement in OSNR penalty compared to that of conventional lasers. For the same OSNR penalty, the optically-filtered laser permits greater parallelization, e.g., increase from 16 to 20, to reduce the hardware processing rate from 312.5 to 250 MHz.

In my second project (Chapter 4), we pointed out, for the first time, that the standard frequency offset compensation algorithm could be the cause of real-time phase tracking failure for 64-QAM in the presence of sine tones in lasers below 1 MHz. We demonstrated experimentally the impact of sinusoidal laser phase noise on phase recovery in the presence of parallelization and pipelining delay. We experimentally investigated, for the first time, the

range of the sine tone amplitude and frequency that lead to real-time phase tracking failure.

1.2.3 Problem Statement for Chapter 6 and Chapter 7

In digital coherent receivers, for the last decade polarization demultiplexing (PolDemux) is performed using 2-by-2 multiple-in multiple-out finite-impulse-response filters (MIMO-FIRs). This technique finds the Jones matrix of the inverse of an optical channel. Conventional blind algorithms, such as the constant modulus algorithm (CMA), used to update the FIR filter coefficients suffer from singularity and convergence problems during FIR tap initialization. Since the emergence of CMA-based MIMO-FIRs, the solution used in industry for singularity and convergence problems is using training sequences to initialize the FIR taps. However, training sequences reduce transmission capacity and burden burst-mode receivers for packet-switching-based metro networks. As a separate issue, the fractional-spaced FIR filters require tens of complex multiplications per symbol, leading to high power consumption. This complexity requirement is due as much to equalization of the limited frequency response of components as to polarization demultiplexing.

1.2.4 Contributions for Chapter 6 and Chapter 7

In my third project (Chapter 6), we propose an innovative technique that corrects singularity and convergence problems, while actually reducing the number of DSP operations per symbol. A low complexity algorithm is used to produce a crude estimate of the received state-of-polarization (SOP) before PolDemux. The accuracy of the estimate was sufficient to derotate the received signal to a benign residual SOP where convergence of PolDemux is rapid. The derotation also allowed us to relax the number of MIMO-FIR cross taps, as those taps are predominantly for demultiplexing, while the through taps performed the role of equalization.

In my fourth project (Chapter 7), we improve our SOP estimation approach so that we maintain our previous advantages, while adding additional functionality. Timing phase must be extracted from the received signal before PolDemux. However, timing phase estimation (TPE) fails at high polarization crosstalk and half-symbol walk-off between two polarization channels. Various algorithms can estimate SOP, but involve sophisticated computations at the symbol rate or higher using a substantial numbers of multipliers. We take these techniques and adapt them to this application, reducing complexity and applying the solution simultaneously with MIMO-FIR tap reduction, as before.

Unlike the decade-old conventional approach, we generate our own channel-state information (CSI), i.e., polarization information in Stokes space. Our algorithm is based on a low-complexity non-data/decision-aided discrete-time extended Kalman filter (EKF) that minimizes a single resultant Stokes parameter (s_1), updated only at every ASIC (application-specific integrated circuits) clock. Thus we track the inverse Jones matrix of the optical channel, and perform an SOP pre-rotation before MIMO-FIRs. This SOP pre-rotation can

coarsely reject the polarization crosstalk before the subsequent PolDemux. It offers several advantages over conventional MIMO-FIR approaches. Firstly, only one instead of three Stokes parameters is needed, greatly reducing the computational effort for estimating SOP compared to other Stokes-space approaches. Secondly, it avoids the problem of long MIMO convergence and singularity for certain SOPs. Thirdly, thanks to the reduced polarization coupling, we are at liberty to reduce the number of MIMO-FIR cross-taps (responsible for compensating crosstalk), leading to a significant reduction in number of MIMO-FIR multipliers. The second and third advantages are derivatives of our previous work, however, use of the EKF reduces the memory depth for SOP estimation. Finally, the SOP pre-rotation also brings the benefit of restoring clock tones for TPE even before PolDemux.

1.3 Organization of thesis

This thesis covers mainly two research areas in digital coherent receivers. The first research area is about phase recovery. Background on phase recovery and a literature review are provided in Chapter 2, while our new contributions are presented in Chapters 3 and 4. The other research area is about polarization demultiplexing. The background of polarization demultiplexing is provided in Chapter 5, while our new contributions are presented in Chapters 6 and 7.

In Chapter 2, the basics of three best-known carrier phase recovery algorithms in digital coherent receivers will be given. The principles of parallelization and pipelining, implementation of phase recovery, the reduced tracking bandwidth and the power spectral density of laser frequency noise are important for the subsequent two chapters.

In Chapter 3, TeraXion's novel narrow-linewidth laser sources having low linewidth using optical filtering are introduced. The principle of laser characterization using coherent setup and the experimental results are presented, followed by the results of system performance in a single-carrier single-polarization 5 Gbaud 64 QAM system.

In Chapter 4, the failure of parallel phase tracking for 64 QAM using laser sources having sinusoidal phase modulation will be illustrated. The result of system performance in a single-carrier single-polarization 5 Gbaud 64 QAM system will be presented, supported by simulation results.

In Chapter 5, a fiber model and the conventional approach for digital polarization demultiplexing are given. Roles and problems of the digital approach are explained in detail, which serve as the motivation for the subsequent two chapters.

In Chapter 6, the principle of blind SOP search for obtaining the polarization information of an optical channel, and the principle of polarization pre-rotation to coarsely reject the polarization crosstalk, will be explained, followed by experimental results of bit-error-rate

performance

In Chapter 7, the extension of Chapter 6, using a low-complexity discrete-time extended Kalman filter to track the polarization effect of an optical channel, will be explained, followed by experimental results of bit-error-rate performance.

Finally, Chapter 8 gives a summary of the conclusions drawn from all the experimental results in this thesis.

Chapitre 2

Phase Recovery Algorithms

In coherent communications, we modulate electrical data (at gigabauds) on a laser source centered within the telecommunication C band (193 THz). At the receiver side, another laser source having the same nominal frequency is used to bring the bandpass optical signal back to the baseband electrical signal. Both laser sources contain phase noise that rotate the complex signal on the I-Q plane, introducing errors upon detection. Phase recovery refers to the process of estimating and compensating the phase noise corrupting received signals.

In the present chapter, background knowledge for Chapter 3 and 4 will be given. A brief review of laser phase noise statistics and phase-noise generation in simulation will be given in Appendix B. In this chapter, we will give an introduction of the basic components of a back-to-back, single-polarization coherent communication systems. A simple formulation of the phase-noise problem will follow. We will cover three best-known algorithms for phase recovery. The concept of parallelization and pipelining will be introduced to underscore the implementation form of maximum-likelihood phase estimation considered in Chapter 3 and 4. Finally, we will classify three operation regions of phase recovery from the viewpoint of frequency-noise power spectral density (FN-PSD) of laser sources.

2.1 DSP Blocks for Single-Polarization Coherent System

We begin with a simple single-polarization, single-carrier coherent system [56, 32] in Fig.2.1a, and its equivalent mathematical model is shown in Fig. 2.1b. The elements of coherent systems are as follows :

1. **I-Q Modulator** : An in phase-quadrature-phase (IQ) modulator is a device for electrical-to-optical conversion. The baseband electrical signals (real and imaginary parts of the complex signal) modulate the amplitude and phase of a continuous-wave, highly coherent laser source via the I-Q modulator [1]. The modulator output is a bandpass optical signal

1. An I-Q modulator consists of two nested dual-drive Mach-Zehner modulators (MZM) in push-pull opera-

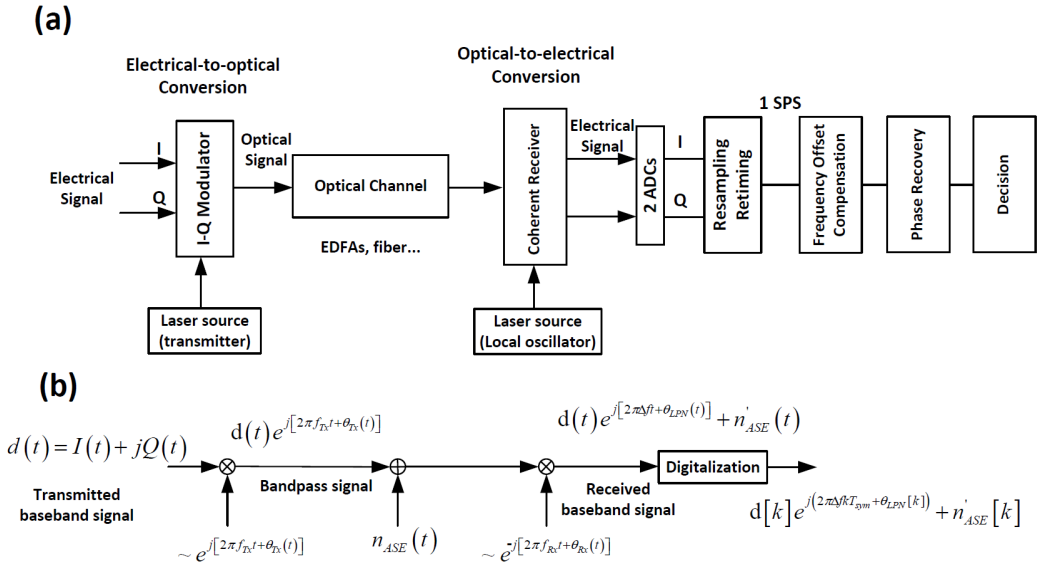

FIGURE 2.1 – (a) Element of a single-polarization single-carrier coherent communication systems. (b) Equivalent mathematical model.

centered at 193 THz. This is equivalent to multiplying our complex baseband signal, $d(t)$, with $e^{j2\pi f_{\mathrm{TX}}t+j\theta_{TX}(t)}$ in Fig. 2.1b, where f_{TX} and θ_{TX} are the frequency and phase noise of the transmit laser, respectively, j is defined as $\sqrt{-1}$, known as the up-conversion in communication theory.

2. **Optical Channel** : The optical channel usually includes distortion coming from the fiber channels such as chromatic dispersion, polarization mode dispersion, fiber nonlinear effect and amplified spontaneous emission (ASE) noise from erbium doped fiber amplifiers (EDFAs). In Chapter 3 and 4, we will only consider a back-to-back system, i.e., only ASE noise, in order to investigate the impact of phase noise and residual frequency offset on 64 QAM. This ASE noise can be described as the additive white Gaussian noise (AWGN), denoted as $n_{ASE}(t)$ in Fig. 2.1b.

3. **Coherent Receiver** : A coherent receiver consists of an optical hybrid and balanced photodetectors. The optical hybrid is a device to assist the separation of the polarizations and real and imaginary parts of the incoming complex signal via exploitation of a local oscillator. Optical-to-electrical conversion is a square-law process [2] in photodetectors [3]. A coherent receiver mixes both signal and local oscillator together, bringing the passband optical signal back to the baseband electrical signal. The center frequencies of both laser sources at transmitter and receiver should be close to each other. Otherwise, the

tion ; the electrical signal of I or Q is applied to each MZM for modulating the laser source. A 90-degree phase shift on either I or Q branch makes the optical fields from the two MZMs become in quadrature. Subsequent optical combination simply results in an optical field with complex data modulation.

2. A square-law device mixes two signals at carrier frequencies together, resulting in a double frequency term out of the bandwidth of the photodetector, as well as a term close to zero-frequency (DC).

3. An optical power is converted into a photocurrent by a photodetector.

electrical signal cannot be baseband, i.e., the detected signal will be partially out of the bandwidth of the photodetector, leading to distortion. Coherent detection is equivalent to multiplying the complex bandpass received signal with $e^{-j\left[2\pi f_{\mathrm{RX}}t+\theta_{\mathrm{RX}}(t)\right]}$ in Fig. 2.1b, where f_{RX} and θ_{RX} are the frequency and phase noise of the transmit laser, known as the down-conversion in communication theory. The downconverted signal becomes $d(t)e^{-j\left[2\pi\Delta ft+\theta_{LPN}(t)\right]}$ is the frequency match between the transmit and receive lasers, where $\Delta f = f_{\mathrm{TX}} - f_{\mathrm{RX}}$, $\theta_{LPN}(t) = \theta_{\mathrm{TX}}(t) - \theta_{\mathrm{RX}}(t)$. The mismatch Δf between two laser frequency must be eliminated by frequency offset estimation, while the random laser phase noise due to θ_{TX} and θ_{RX} must be compensated by phase recovery. The ASE noise $n'_{ASE}(t)$ is a randomly-rotated version of $n_{ASE}(t)$.

4. **Analog-to-digital conversion (ADC)** : The ADCs perform sampling to digitize the analog electrical baseband signal from the physical world, at a sampling rate defined by system designers. To facilitate the analysis in Chapter 3 and 4, here we only start with one sample per second (SPS) [4]. The digitized signal contains the quantization error. Due to the limited bandwidth of electrical components, the received signals are distorted by intersymbol interference. The digitized signal is shown in Fig. 2.1b, where k is the discrete time index, and T_{sym} is the symbol duration.

5. **Resampling and retiming** : This is the starting point of our DSP for a back-to-back coherent system. The digitized signal shown in Fig. 2.1b is an ideal one. First, the clock frequency between transmitters and ADC may not be the same, but throughout my dissertation this clock mismatch will not be considered because in our experiments we synchronized the transmitter with the receiver using the same clock source. Second, ADCs cannot recognize the exact timing for sampling even in the absence of the clock mismatch. When the ADC sampling rate is different from the symbol rate, one has to first apply interpolation for upsampling the ADC samples, and then downsample to match the symbol rate. This process is referred as resampling. Retiming is simply to choose a sample within a symbol duration that represents the center point of a symbol duration.

6. **Frequency Offset Compensation** : The frequency offset between transmit and receive lasers deterministically rotates the signal constellation. This is our focus in Chapter 4.

7. **Phase recovery** : Laser phase noise rotates the constellation randomly. Phase noise estimation algorithms are required. This is our focus in Chapter 3.

8. **Symbol Detection using Hard decision** : Only hard decision is considered in this thesis. This is a process to decide what symbols we receive, based on setting thresholds between different symbols.

Although there exists some algorithms for joint estimation of frequency offset and phase noise, we will adhere to standard practice in industry that assumes frequency offset is small (within

4. It is because the DSP for phase recovery and detection only requires one sample per second.

50 MHz) after the frequency offset compensation, meaning that the electric field (complex signal) is baseband. It is because the equalization of channel impairment and the intersymbol interference introduced by the limited receiver bandwidth must be performed before the phase estimation. Otherwise, the constellation positions will deviate from the standard ones [32].

2.1.1 Signal Model

We start our problem statement by assuming a perfect channel equalization, meaning that the received signal is corrupted by ASE noise only. The received signal r_k after perfect channel equalization or back-to-back operation is modeled as :

$$r_k = d_k e^{j\theta_k} + n_k, \tag{2.1}$$

where d_k is a complex-valued symbol transmitted at the k^{th} symbol period, θ_k is the real-valued carrier phase, and n_k is the additive white Gaussian noise (AWGN). In terms of detection and estimation theory, d_k and θ_k are parameters, while r_k is the observation (received signal) corrupted by the complex-valued measurement noise n_k.

Our goal is to obtain a good estimate $\hat{d}_k$ of data, based on the observation r_k. However, without a good phase estimate $\hat{\theta}_k$, the estimated data may not be correct. We have a problem of joint estimation [5] of θ_k and d_k [23]. In fact, as we see later, we will using three techniques to remove the effect of data, so that we can first estimate θ_k. Subsequently, symbol detection using hard decision allows us to estimate the transmitted data.

2.2 Estimation-based Method

Since the parameters θ_k are random, this parameter estimation belongs to the class of **Bayesian estimation** [24, 47] [6], i.e., using the conditional probability density function (pdf) of the transmitted symbols d_k and laser phase noise θ_k given the observation r_k, $p(d_k, \theta_k | r_k)$, which is also called the posterior pdf [24].A practical Bayesian approach is to use the mode (location of the maximum) of the posterior pdf, termed the maximum a posteriori (MAP) estimation, i.e., to construct an estimator that can maximize the posterior pdf, $p(d_k, \theta_k | r_k)$. Using Bayes' theorem,

$$p(d_k, \theta_k | r_k) = \frac{p(r_k | d_k, \theta_k) p(d_k, \theta_k)}{p(r_k)}, \tag{2.2}$$

where the denominator of (2.2) is the unconditioned density of observation. Thus, the MAP estimate can be obtained without the computation of $p(r_k)$ since this term does not affect the maximization over θ. The random data d_k is uniformly distributed and is independent

5. Since both data and laser phase noise are time-varying, estimation of time-varying parameters is conventionally defined as signal tracking in the field of communication theory [47]. Moreover, since the data is M-QAM signals involving discrete levels or candidates (composite hypothesis), the estimation of data belongs to detection theory[47]. However, in this thesis, we use the term estimation for simplicity.

6. Bayesian estimation theory can be found in Chapter 4 in [47] or in Chapter 10 in [24]

of θ_k. We have no prior information about θ_k and assume that θ_k is uniformly distributed) [7]. The problem becomes the maximization of the likelihood function $p(r_k|\, d_k, \theta_k)$ (ML) over the unknown d_k and unknown θ_k. Since the prior information about the parameter θ_k is not used, this belongs to the class of **non-random parameter estimation** [47]. For the same reason, the following three best-known phase recovery algorithms are all non-random parameter estimation / detection.

2.2.1 Decision-Directed Maximum Likelihood Algorithm

The conditional probability of the received signal given the transmitted signal and the known laser phase noise (called the likelihood function) is given as

$$p(r_k|\, d_k, \theta_k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\left|r_k - d_k e^{j\theta_k}\right|^2}{2\sigma^2}\right], \tag{2.3}$$

where σ is the standard deviation of the complex AWGN, n_k. Maximizing the above likelihood function is equivalent to maximizing the log-likelihood function, [8] $\ln p$, or maximizing the term $\mathrm{Re}(r_k^* d_k e^{j\theta_k})$, where Re denotes the real part of a complex quantity. To estimate θ_k, data should be estimated or removed before phase estimation. This process is called called "**data removal**". Data removal seems impossible without the phase information, but can be performed by using a training sequence to initialize any algorithm. A more practical approach, which avoids training sequence, is to use blind phase search (please see Section 2.2.2) to generate an initial phase, enabling the subsequent decision, $\hat{d}_k$, such that the data can be estimated "in advance". Since a decision is required to aid the phase estimation, it is also called decision-directed algorithm (DD).

We maximize the term $\mathrm{Re}(r_k^* \hat{d}_k e^{j\theta_k})$ over the $\hat{\theta}_k$. Using (2.1), for sufficiently high SNR such that correct decisions are made, the phase estimate becomes

$$\hat{\theta}_k = \max_{\hat{\theta}_k} \mathrm{Re}[(d_k e^{j\theta_k} + n_k)^* \hat{d}_k e^{j\hat{\theta}_k}] = \max_{\hat{\theta}_k} \mathrm{Re}(P_k e^{j(\hat{\theta}_k - \theta_k)} + n'_k), \tag{2.4}$$

where P_k is the M-QAM signal power, n'_k is the rotated AWGN having the same variance as n_k. Equation (2.4) is a nonlinear function and it is difficult to obtain a phase estimate in analytical form.

To realize the maximum likelihood estimation, we follow the flow diagram shown in Fig. 2.2. Ideally, one wishes to de-rotate the received signal r_k by the true phase estimated by (2.4), i.e. multiply r_k in (2.1) by $e^{-j\theta_k}$. To do this, we construct a phasor V :

$$V = e^{-\hat{\theta}_k} = \sum_{i=k}^{k+L-1} \frac{r_i^* \hat{d}_i}{\left|r_i^* \hat{d}_i\right|} \tag{2.5}$$

7. In fact, laser phase noise has a known ensemble statistics, which can improve phase estimation.

8. The log-likelihood function : $\ln p(r_k|\, d_k, \theta_k) = -\ln\sqrt{2\pi}\sigma - \frac{\left|r_k - d_k e^{j\theta_k}\right|^2}{2\sigma^2}$, where the constant factors can be ignored as they are independent of parameters.

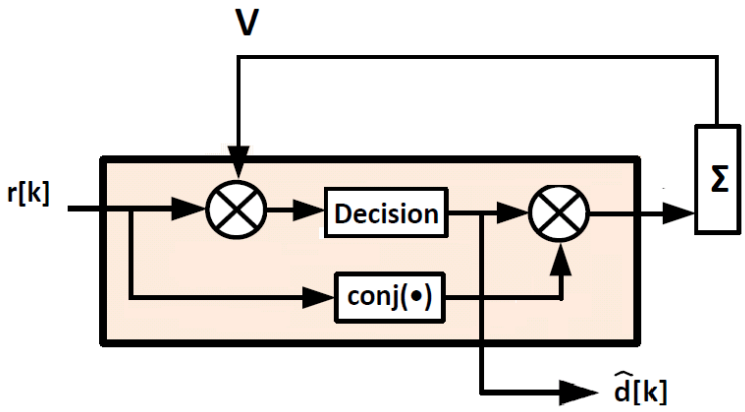

FIGURE 2.2 – Block diagram of decision-directed maximum likelihood phase estimation.

The summation with a block length of L is to suppress the AWGN. For high SNR, V approaches the true $e^{j\theta_k}$. The output of the first multiplier in Fig. 2.2 becomes $d_k e^{j(\theta_k - \hat{\theta}_k)}$. For small phase error, the correct decision can be made, i.e., $\hat{d}_k = d_k$. The estimate $\hat{d}_k$ will be multiplied with the conjugate of r_k to generate the phasor estimate V, which is fed back to the first multiplier.

To initialize the algorithm, we need to know the first phasor estimate by using a training sequence. Another method is to use blind phase search (which is assisted by a known constellation). However, due to ambiguity, the data may have different possible values (inverse polarities in I and Q) that require further processing to determine the polarity of transmitted data. Feedback greatly complicates the algorithm from hardware parallelization for computation.

2.2.2 Feedforward Viterbi-Viterbi Algorithm

Contrary to the decision-directed algorithm, non-data-added (NDA) phase estimation avoid training sequences and decision feedback [23]. The Viterbi-Viterbi algorithm for M-PSK is an example of an NDA algorithm, in which we first remove the data modulation by raising the received M-ary PSK signal, $r_k = \exp[i\theta_{k,MPSK}] + n_k$, where $d_k = 1$ for constant-envelop modulation format, $\theta_{k,MPSK} \in \{n\frac{2\pi}{M}, n = 0, 1, ..., M-1\}$, to the Mth power [65] :

$$r_k^M = \exp(iM\theta_{k,MPSK}) + O(d_k, n_k) + n_k^M = \exp(iM\theta_{k,MPSK}) + q_k, \tag{2.6}$$

where $O(d_k, n_k)$ is the data-AWGN beating terms, $q_k \triangleq O(d_k, n_k) + n_k^M$ [23], subscript k is the discrete-time index. We need to obtain the phase angle of r_k^M.

However, cycle clip occurs because the angle calculated using "arc-tangent" (or the function "angle" in Matlab) only returns the phase angle in the principal region between $-\pi$ and π, which is called phase ambiguity. Therefore, the phase that is supposed to evolve and extend to positive and negative infinity is being wrapped inside this principal region. We can use the

standard Matlab function "unwrap" to unwrap the phase [9], or use the algorithms mentioned in [23, 65]. We assume that q_k is a small additive noise, the unwrapped phase becomes [65]

$$\arg\left\{r_k^M\right\} = M\theta_{k,MPSK} + n'_k, \tag{2.7}$$

where n'_k is an one-dimensional noise To estimate θ_k, one uses a moving average (MA) filter [10] to remove the AWGN effect in coherent receivers, such that

$$\hat{\theta}_{k,MPSK} = \sum_{i=k}^{k+L-1} \frac{1}{M} \arg\left\{r_i^M\right\}, \tag{2.8}$$

where L represents the MA filter length. In practice, for low OSNR (after long-haul transmission), the AWGN is large and causes more of the aforementioned phase ambiguities. To reduce the probability of having more cycle slips, the moving average is performed before the angle function :

$$\hat{\theta}_{k,MPSK} = \frac{1}{M} \arg\left\{\sum_{i=k}^{k+L-1} r_i^M\right\}, \tag{2.9}$$

Finally, we apply the phase estimate to the complex field using $e^{-j\hat{\theta}_{k,MPSK}}$ and to the received signal r_k to remove the phase noise. The estimated data becomes

$$\hat{d}_k = r_k e^{-j\hat{\theta}_{k,MPSK}} = d_k e^{j\theta_{k,MPSK} - j\hat{\theta}_{k,MPSK}} + n_k e^{-j\hat{\theta}_{k,MPSK}}. \tag{2.10}$$

For perfect phase estimation, the estimated data is equal to the transmitted data.

Feedforward Viterbi-Viterbi Algorithm for MQAM

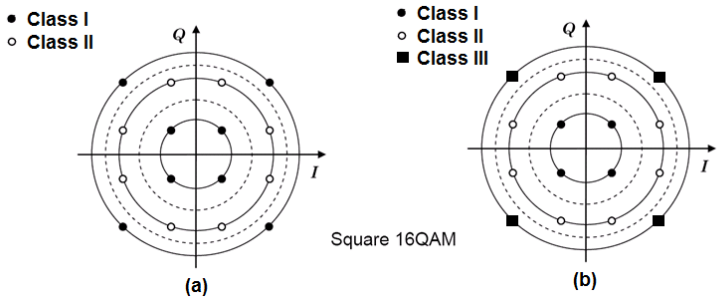

FIGURE 2.3 – Symbols classification used in Viterbi-Viterbi algorithm (modified for MQAM) in (a) Seimetz's work [52] and (b) Gao et al.'s work [13] .

For M-QAM, the M-th power approach is not effective at every constellation point. For 16-QAM, Seimetz [52] classified the constellation points into two subgroups (Class I and II)

9. For implementation, phase recovery is a symbol-wise operation, and therefore the cycle-slip detection and correction must be done symbol-wise, enhancing the hardware complexity due to parallelization.

10. While a moving average filter is suboptimal (the Wiener filter is optimal [65]), it is used for simplicity.

shown in Fig. 2.3. Only the symbols in the subgroup Class I, i.e., $\frac{\pi}{4}+n\frac{\pi}{2}$, $n=\{0,1,2,3\}$, will be chosen for phase estimation, and the estimated phase noise for square M-QAM becomes

$$\hat{\theta}^{I}_{k,MQAM}=\frac{1}{4}\arg\left\{\sum_{i=k}^{k+N_I-1}\frac{r^4_{i,I}}{\left|r^4_{i,I}\right|}\right\}, \tag{2.11}$$

where $\hat{\theta}^{I}_{k,MQAM}$ is the estimated phase using the received symbols in Class I, $r_{k,ClassI}$ is the received signal falling into the subgroup Class I, N_I is the number of Class I symbols used to generate $\hat{\theta}^{I}_{k,MQAM}$. Since only a subgroup of received signals is used, the algorithm is sensitive to laser phase noise. Gao et al. [13] proposed a modified Mth power approach, by classifying the inner-ring symbols and the outer-ring symbols as Class I and Class III, respectively, and use both classes for phase estimation :

$$\hat{\theta}^{I,III}_{k,MQAM}=\frac{1}{4}\arg\left\{\sum_{i=k}^{k+N_I-1}\frac{r^4_{k,I}}{\left|r^4_{k,I}\right|}+W_1\sum_{i=k}^{k+N_{III}-1}\frac{r^4_{k,III}}{\left|r^4_{k,III}\right|}\right\}, \tag{2.12}$$

where $\hat{\theta}^{I,III}_{k,MQAM}$ is the estimated phase using the received symbols in both Class I and Class III, $r_{k,III}$ is the received signal falling into the subgroup Class III, N_{III} is the number of Class III symbols used to generate $\hat{\theta}^{I}_{k,MQAM}$, W_1 is a weighting coefficient which takes into account the fact that symbols in the outer ring have a higher OSNR and hence carrier phase estimates from these symbols should be better than those from the inner ring. Moreover, Fatadin et al. [9] shifted the middle-ring symbols by $\pm 19.5^o$ in order to fully make use of all symbols for phase estimation :

$$\hat{\theta}_{k,MQAM}=\hat{\theta}^{I,III}_{k,MQAM}+W_2\hat{\theta}^{II}_{k,MQAM}, \tag{2.13}$$

where $\hat{\theta}^{II}_{k,MQAM}$ is the estimated phase corresponding to the middle ring, W_2 is another weight coefficient for the estimated phase using Class II received symbols.

2.3 Detection-based Method

The concept of the blind-phase search (BPS) [44] is to estimate the laser phase noise of the received signal, by making a choice between different (discrete, countable) phase candidates that can minimize the defined cost function. As the parameters are discrete, not a continuum of possible values, this problem belongs to the **composite hypothesis** testing in **detection theory** [47].

2.3.1 Feedforward Blind-Phase Search (BPS)

As shown in Fig. 2.4, the received signal after sampling, Y_k, is first rotated by B test phase angles φ_b by multiplying Y_k with B phasors terms $e^{j\varphi_b}$, where

$$\varphi_b=\frac{b}{B}\cdot\frac{\pi}{2}, b=\{0,1,...,B-1\}. \tag{2.14}$$

Note that, since square M-QAM has a $\pi/2$-rotational symmetry, the maximum range of the test phase angles is $\pi/2$, meaning that these test phase angles cover only one quadrant on the complex plane. As the laser phase noise rotates the constellation of the transmitted square

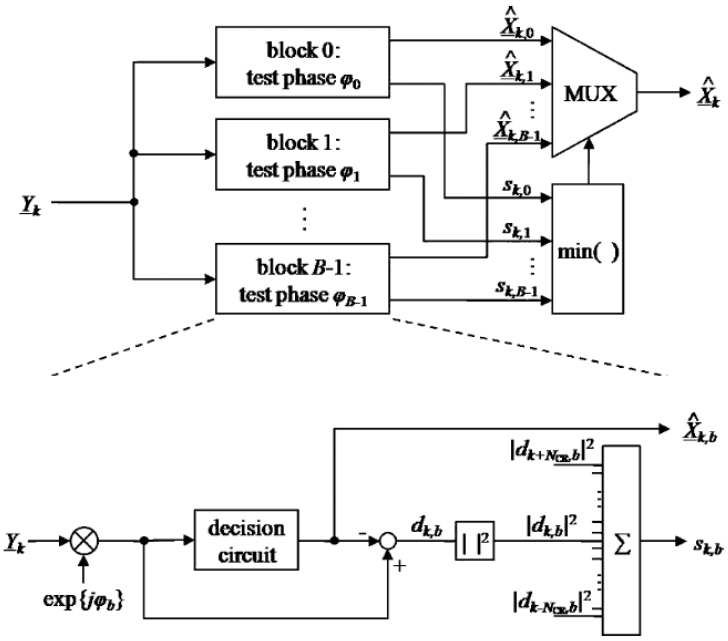

FIGURE 2.4 – Block diagrams of feedforward blind phase search, reproduced from Fig. 4 in [44].

M-QAM symbols, the corrupted squared M-QAM symbols are no longer at their original positions. The interesting point of this blind phase search is that it makes use of the original constellation shape of squared M-QAM to undergo **data-removal**. The followings show the step of BPS :

1. **De-rotation using each test phase** : De-rotate each received symbol by test phase angles shown in Fig. 2.5.

2. **Calculation of cost function** : Calculate the cost function of the B de-rotated received symbol. The cost function is defined as the Euclidean distance, $d_{k,b}$, between the de-rotated receiver symbol and the original position of square M-QAM symbols, where k is the time index of the received symbols.

$$|d_{k,b}|^2 = \left| Y_k e^{j\varphi_b} - \hat{X}_{k,b} \right|^2, \tag{2.15}$$

   where Y_k is the sampled received signal, $\hat{X}_{k,b}$ is the output of the decision device, which is equal to one of the M-QAM symbol whose position is the closest to the rotated received symbol.

3. **Moving average** : To reduce the impact of the ASE noise, one may use a moving average filter to estimate the Euclidean distances taken within a block of $2N+1$ symbols, shown in the bottom part of Fig. 2.4.
4. **Searching optimal test phase** : There is only one test phase closest to the true phase, resulting in a minimum average Euclidean distance.
5. **Compensation** : The optimal phase is used to de-rotated the received symbol.

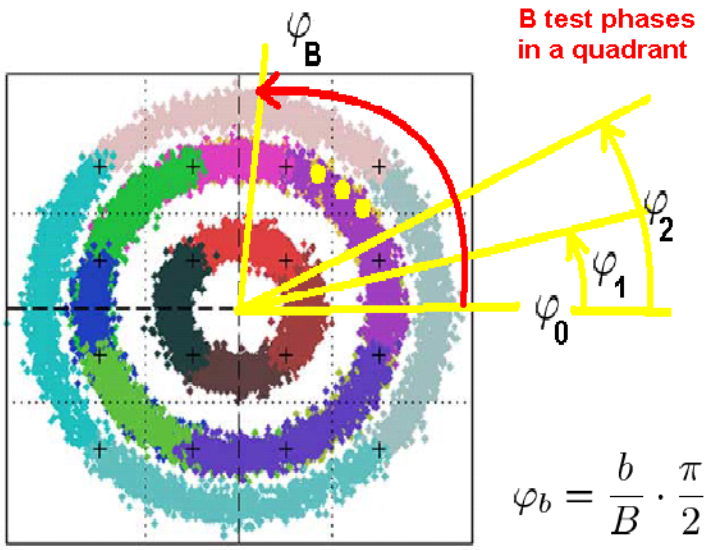

FIGURE 2.5 – B test phases in a quadrant. Background : 16-QAM with uncompensated phase noise [44].

This blind phase search does not require phase unwrapping, since the estimation is done on the complex plane (directly on the electric field). However, because of the $\pi/2$-rotational symmetry of square M-QAM, 4-fold ambiguity leads to a new problem : for example, when the phase noise passes from the first quadrant to the second quadrant, it will be misinterpreted as the phase angle in the first quadrant (the current phase value minus $\pi/2$). It is because, as aforementioned, the range of the test phase angles for PMS is only from zero to $\pi/2$ (covering only the first quandrant).

Due to the 4-fold ambiguity of the recovered phase in squared M-QAM, differential encoding will be used to encode the first two bits of each symbol. For details, please refer to Section III in [44].

2.4 Parallelization and Pipelining

There are two practical problems in the implementation of DSP for optical communication. First, power consumption of CMOS chips increases linearly with clock frequency [[42], (3.10)],

and therefore increases with data rate ; DSP at lower clock speed provides power savings. Second, the currently available ASICs and reconfigurable FPGA cannot process at Gbaud rates. Parallel processing and pipelining are required, which tremendously increases the effective linewidth of lasers.

For parallel processing, the serial data is acquired in a time-interleaved manner into several parallel channels having duplicate processing hardware [42]. This allows processing multiple data in parallel in a clock period. With **time-interleaved parallelization**, the received symbols at Gbaud rates (symbol period T_{sym}) are demultiplexed into P sub-GHz channels for DSP. For example, in Fig. 2.6, during the first round of time interleaving P symbols, $r[k]$, $r[k-1]$, ..., $r[k-P+1]$ arrive at the input of the first, second, ..., P^{th} channel, respectively. Next, another P symbols, $r[k+P]$, $r[k+P-1]$, ..., $r[k+1]$ follow, and the process will repeat. Thus, the i^{th} channel receives time-interleaved symbols $r[k-i+1]$, $r[k-P-i+1]$, $r[k-2P-i+1]$, and so on, where i is the index of parallelization channels, taken from 1 to P, where P is the number of parallelization. Two adjacent symbols in each channel have a time separation of $P \times T_{sym}$, instead of T_{sym} in serial processing. As a result, each channel perceives laser phase noise having a variance of $2\pi(P\Delta\nu)T_{sym}$, i.e., time-interleaved parallelization increases the effective laser linewidth by a factor of P.

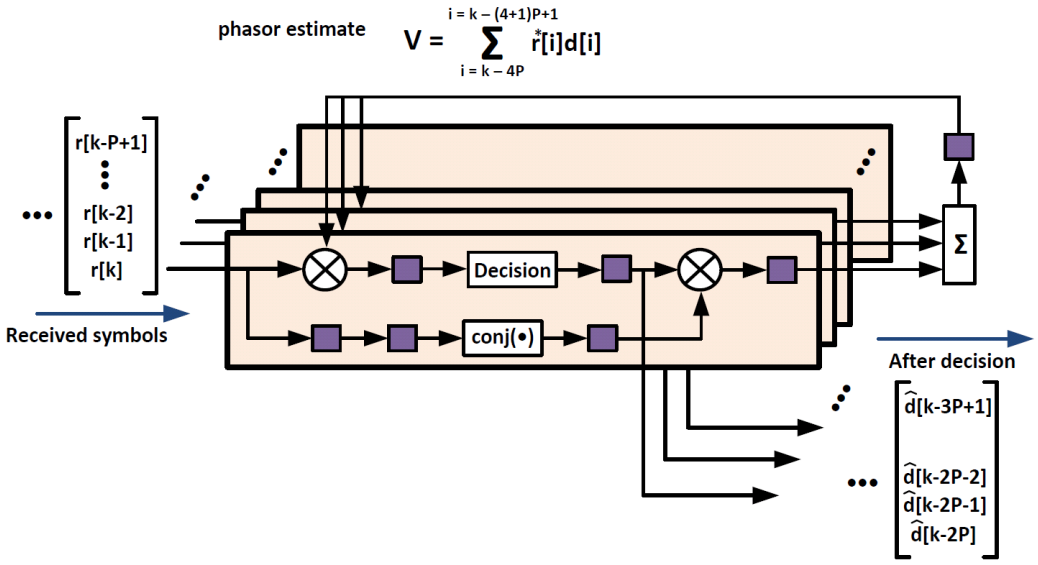

FIGURE 2.6 – Implementation of decision-direct maximum likelihood estimation for a time-interleaved structure [75].

Even though ASIC design allows higher processing speed than FPGAs by circuit optimization, multipliers and lookup tables are unavoidable sources of latency. The throughput of an

algorithm simply cannot keep up with the input data rate (sampling rate) at each parallelization channel. Pipelining must be introduced in order to increase the effective sampling rate [42]. For example, an algorithm usually consists of multipliers, slicers and lookup tables that cannot be implemented within one clock cycle. Thus pipelining is required, in which delays (flip-flops [44] or so-called pipelining registers [42]) are added between individual operations requiring a clock cycle or less ; the immediate result of each operation is stored in registers for subsequent manipulation.

2.5 Implementation of DDMLE

Decision-directed phase recovery (DD-PR) is the preferred solution for high baud rate, higher-density constellation systems. The serial DD-PR intrinsically contains one symbol delay on the feedback path. To realize parallel DD-PR in ASIC or FPGA, however, the feedback delay is increased. To illustrate the effect of feedback delay in parallel DD-PR, we select the decision-directed maximum likelihood estimation (DD-MLE).

For serial tracking, DD-MLE can be linearized as the following first-order discrete loop equation [[15], (6)],

$$\hat{\theta}_{k+1} = \hat{\theta}_k + U, \tag{2.16}$$

where $\hat{\theta}_k$ is the phase estimate, and U is the phase increment ; U is a function of the parameters of the moving average filter, laser phase noise variance and SNR. For parallel tracking with feedback delay, as shown in Appendix D, (2.16) can be modified as

$$\hat{\theta}_{k-i+1} = \hat{\theta}_{k-D-i+1} + f\Big(\{\underline{\theta}, \hat{\theta}_{k-D-i+1}, \underline{d}, \underline{n}_1^*\}\Big), \tag{2.17}$$

where

$$\underline{\theta} = [\theta]_{k-D}^{m=k-D-P+1}, \tag{2.18}$$

$$\underline{d} = [d_m]_{k-D}^{m=k-D-P+1}, \tag{2.19}$$

$$\underline{n}_1^* = [n_{1,m}^*]_{k-D}^{m=k-D-P+1} \tag{2.20}$$

where i is the parallelization channel index taking from 1 to P, $D = P \times d$, d is the number of pipelining registers in the closed loop, $\underline{\theta}$, $\underline{d}$ and $\underline{n}_1^*$ are vectors of true phases, transmitted symbols and the AWGN, respectively, taking from $k - D - P + 1$ to $k - D$. The second term of right-hand side of (2.17) is the increment of the discrete loop equation, which is a function of phase tracking error, transmitted symbols and the AWGN seen in each rail taking from $k - D - P + 1$ to $k - D$. For higher parallelization levels (required to process Gbaud rates in optical communications, e.g. P = 20 for 5 Gbaud signal and 250 MHz ASIC speed), the feedback delay of D symbols results in much higher phase error than that of serial tracking. For $d > 1$, the effect of the feedback delay dominates that of the moving average filter, and

therefore the introduction of feedback delay equivalently reduces loop bandwidth and phase margin [15].

Fig. 2.6 shows that the implementation of DD-MLE has four pipelining registers in the closed loop [75], i.e., $d = 4$. The effect of the loop delay can be conceptually illustrated as in Fig. 2.7. The estimated phase is generated using $(k\text{-}5P{+}1)^{th}$ to $(k\text{-}4P)^{th}$ symbols. The compensation is applied to the $(k\text{-}P{+}1)^{th}$ to k^{th} received symbols, introducing a lag of approximately $D = 4 \times P \times T_{sym}$ seconds. Delayed phase noise compensation caused by pipelining registers on feedback path further increases phase error variance to approximately $2\pi(4P\Delta\nu)T_{sym}$, i.e., the effective laser linewidth further increases in the DD-MLE case by $4P$. Pipelining delays are detrimental to real-time feedback-delayed phase tracking at Gbaud-rate optical communications.

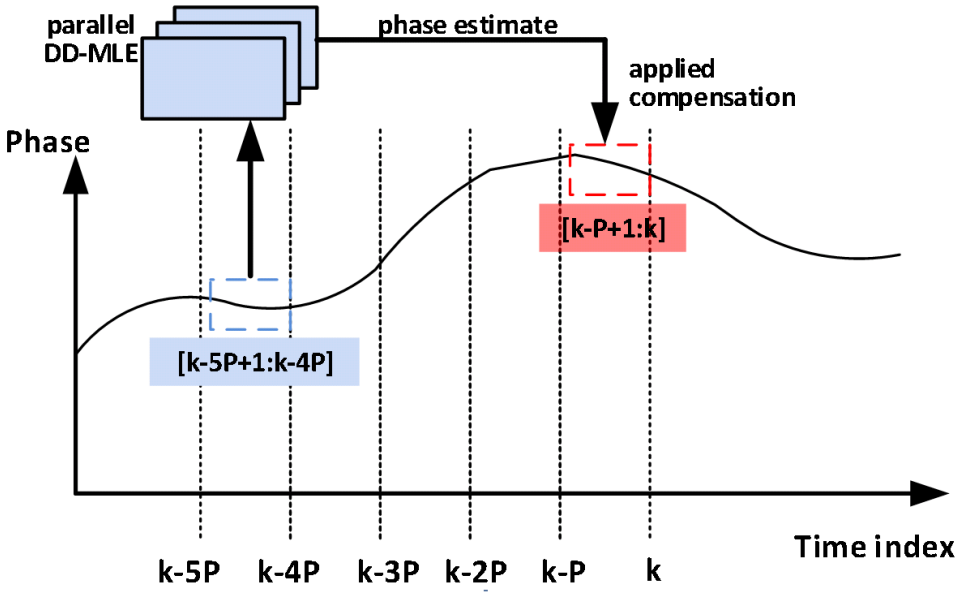

FIGURE 2.7 – Equivalence of implementation of DD-MLE with feedback delay.

2.6 Frequency Noise Power Spectrum Density

In this section, we will introduce three frequency ranges of frequency noise power spectral density (FN-PSD), which will be useful in Chapters 3 and 4. We will later see that phase recovery depends on the FN-PSD level rather than laser linewidth. Thus, one may use the FN-PSD as a figure of merit instead of laser linewidth to understand the impact of the laser sources on the phase tracking. As shown in Fig. 2.8, we can divide the FN-PSD into three different regions based on offline processing, parallel processing and non-ideal frequency offset compensation :

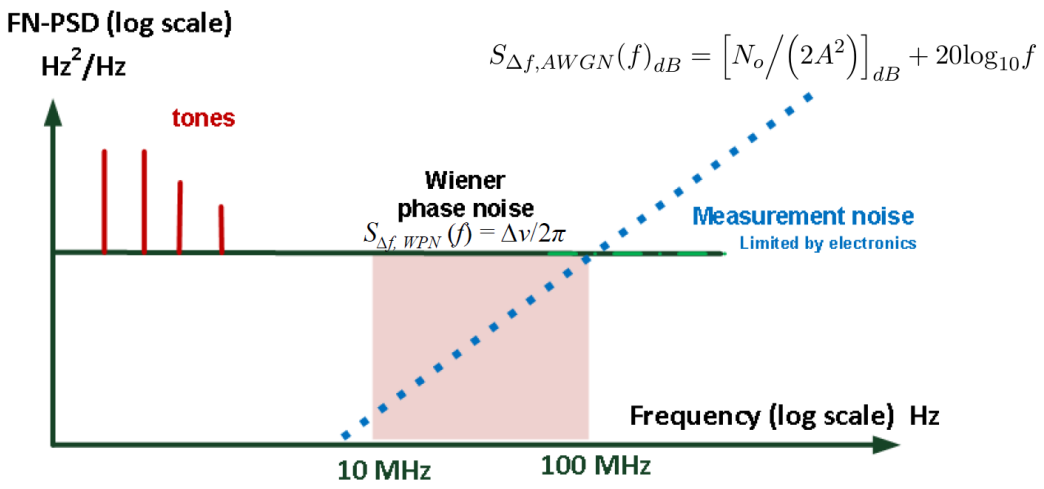

FIGURE 2.8 – Laser Frequency noise power spectral density.

- **Above 100 MHz** : This region is pertinent to all kinds of algorithms for theoretical analysis or offline processing that assume the absence of feedback delay in phase recovery algorithms.
- **10 MHz-100 MHz** : This region affects the decision-directed phase recovery in the presence of feedback delay for QAM signals. We will devote Chapter 3 to explain how TeraXion's narrow-linewidth lasers can outperform the conventional laser sources due to the frequency noise suppression in this frequency range.
- **Below 1 MHz** : The laser frequency noise in this region does not directly affect the phase tracking. However, any frequency tones in this region, depending on the amplitude or frequency, could result in significant residual frequency offset using the conventional frequency offset compensation techniques. The large residual frequency offset may result in phase tracking failure caused by the reduced bandwidth due to feedback delay. We will devote Chapter 4 to this problem.

2.7 Summary

In this chapter, we have revisited the three conventional phase recovery algorithms : Viterbi-Viterbi phase estimation, decision-directed maximum likelihood estimation and blind phase search.

For commercial DP-QPSK coherent systems, Viterbi-Viterbi phase estimation is used because of its simplicity and feedforward structure. For DP-16QAM, BPS is a good candidate. However, BPS suffers from its high computational effort, as BPS calculates the error functions

for all the test phase candidates. Therefore, the modified Viterbi-Viterbi phase estimation is preferred for 16-QAM in order to avoid feedback delay.

For 64-QAM or higher, the modified Viterbi-Viterbi phase estimation shows extra optical signal-to-noise ratio (OSNR) penalty due to amplitude discrimination and its stringent laser linewidth requirement. The BPS requires tremendous computational effort to provide the necessary phase resolution. DD-MLE requires the least computational effort and does not require amplitude discrimination, and this is typically used for 64-QAM experimental demonstrations with offline DSP [18, 43, 10, 76]. Nevertheless, feedback delay remains an obstacle to implement DD-PR in real-time [76].

Chapitre 3

Optically Filtered Laser

In this chapter, we will examine the performance of TeraXion's narrow-linewidth laser source with a decision-directed phase recovery algorithm as discussed in Section 2.5. We will discuss the operation of TeraXion's novel lasers. We will discuss the experimental results of laser characterization, in which we extracted some physical parameters from the laser sources. For background on laser characterization, please refer to Appendix F. We will discuss the relationship between the laser frequency-noise power spectral density, parallelization level and phase tracking bandwidth for 64 QAM using our heuristic approach. We will present our experimental results for the system performance given by TeraXion's lasers when using parallel and pipelined phase tracking in a single-carrier 5 Gbaud [1] 64-QAM back-to-back coherent system.

3.1 Introduction

Implementation of real-time gigabaud optical coherent systems for single-carrier higher-level modulation formats such as 64-QAM depends heavily on phase tracking. For offline digital signal processing, decision-directed phase recovery (seen in Section 2.2.1) is performed at the symbol rate and has the best performance and the least computational effort compared to other best-known algorithms. However, as the decision-directed phase recovery consists of a feedback proportional to parallelization level and pipelining delay, processing at the symbol rate is impractical in current electronics, as discussed in Section 2.4. This leads to the paucity of experiments demonstrating real-time phase tracking for 64- or higher QAM.

In the present chapter, we will experimentally investigate the improvement in system perfor-

1. We chose 5 Gbaud as our symbol rate because of the limitation of our MICRAM, the digital-to-analog converter (DAC) in our laboratory. For 64-QAM generation, both I and Q require a 8-level signal, and thus the DAC should provide at least 8 levels (3 bits) for a desired symbol rate. However, 8-level signals are more susceptable to limited electrical bandwidth, requiring pre-distortion. As a result, 4 samples per symbols were used to generated two 5 Gbaud 8-level electrical signals using our 20Gsa/s, 6 bit waveform generator (AWG) in MICRAM.

mance of TeraXion's narrow-linewidth compared to conventional lasers, when we consider the use of decision-directed maximum likelihood estimation in its implementation form shown in Section 2.5.

Contribution

In this chapter, we limit our scope for a back-to-back, single-polarization [2], single-carrier 5 Gbaud 64-QAM heterodyne coherent system and consider the standard algorithm (parallel and pipelined DD-MLE) for phase tracking. We have made the following contributions :

1. We demonstrate how the laser linewidth is a poor predictor of DSP phase noise tracking efficiency.
2. We suggest the use of frequency noise power spectral density level to describe the impact of laser phase noise for real-time phase tracking with feedback delay, in contrary to nowadays offline experimental demonstrations in which the absence of feedback delay is assumed.
3. We experimentally show, for the first time, that TeraXion's narrow-linewidth lasers gives BER improvement for 64 QAM. The use of a fiber Bragg grating suppresses the frequency noise spectral level of contemporary narrow-linewidth semiconductor lasers such as external cavity lasers (ECLs) or integrated tunable laser assemblies (ITLAs).
4. We found that the frequency noise of laser sources between 10 MHz and 100 MHz affects the real-time decision-directed phase recovery as a result of reduction of tracking bandwidth due to hardware parallelization and pipelining. This result should hold as well for higher symbol rates than those examined, as a higher symbol rate requires higher parallelization in the current CMOS technology.
5. Compared to conventional laser sources, the TeraXion source permits greater parallelization, e.g., increase from 16 to 20, to reduce the hardware processing rate from 312.5 to 250 MHz.

3.2 TeraXion Novel Lasers

The following information about the laser product is provided by TeraXion.

In this project, we examined the performance of TeraXion's Pure Spectrum^{TM}-Tunable Narrow Linewidth Lasers (PS-TNLs) in optical coherent communications. The PS-TNL is based on an Integrated Tunable Laser Assembly (ITLA), with optical filtering of white frequency noise by an ultra-narrowband multichannel fiber Bragg grating (FBG).

2. The experiment in this chapter was performed when the dual-polarization setup in our laboratory was not ready. In any case, this chapter is about optically-filtered phase noise, and one should only consider a single-polarization scenario to rule out the polarization effect.

The optical filter produces 51 narrowband transmission peaks separated by 100 GHz [2]. The interleaved spectral responses present a narrow transmission peak (FWHM of 49 to 77 MHz for all channels) at every 115 GHz from 191.4 to 196.6 THz, so that one peak can be thermally tuned to any frequency within this range. Fig. 3.1 shows the optical spectrum of the FBG

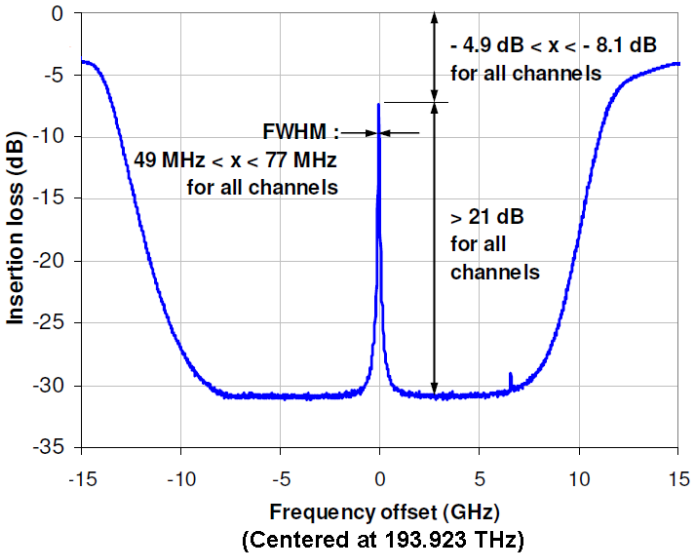

FIGURE 3.1 – Transmission spectrum of the FBG assembly. [2]

assembly including a polarization isolator and a polarizing tap isolator, measured between points A and B on Fig. 3.2. The ITLA has a measured linewidth of 10 kHz.

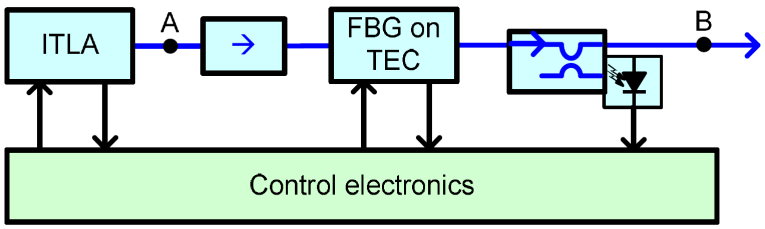

FIGURE 3.2 – Frequency locking schematic. TEC : thermoelectric cooler. [2]

The optical filtering of the laser spectrum is achieved by control of the laser carrier and the narrow transmission peak of the FBG assembly. Fig. 3.2 illustrates the frequency locking mechanism.

A dithering signal modulates the ITLA. The filtered signal is detected and sent to the control

electronics. Correction signals are generated to maintain alignment of the ITLA carrier frequency with a selected transmission peak. Both thermal tuning of the optical filter and ITLA carrier frequency control is applied to correct the misalignment. Eventually slow variations become less detrimental to phase estimation, whereas fast frequency fluctuation associated to the wings of the laser spectrum falling outside of the narrowband transmission peak are suppressed. The optical filtering scheme covers the entire C-band.

3.3 Laser Characterization

3.3.1 Experimental Results

For laser sources having pure Brownian phase noise (or equivalently white frequency noise), the linewidth is defined as the 3-dB width of the Lorentzian field power spectral density (PSD) [7]. To understand the interplay of phase noise reduction with phase tracking algorithms working in a parallel architecture that we have discussed in Chapter 2.4, we must turn to the frequency noise power spectral density (FN-PSD). Brownian phase noise leads to a flat FN-PSD. A cartoon of FN-PSD is given in Fig. 3.3 . The black solid line shows the mid-frequency band (from 10 to 100 MHz) dominated by Brownian phase noise, i.e., $S_{\Delta f,BM} = \Delta\nu/2\pi$, where $\Delta\nu$ is the double of the original laser linewidth (because of the mixing in homodyne detection) ; it is constant over frequency. Conventionally, laser linewidth allows a quick quantification of phase-error variance in coherent communication systems with Brownian phase noise for system performance analysis [40].

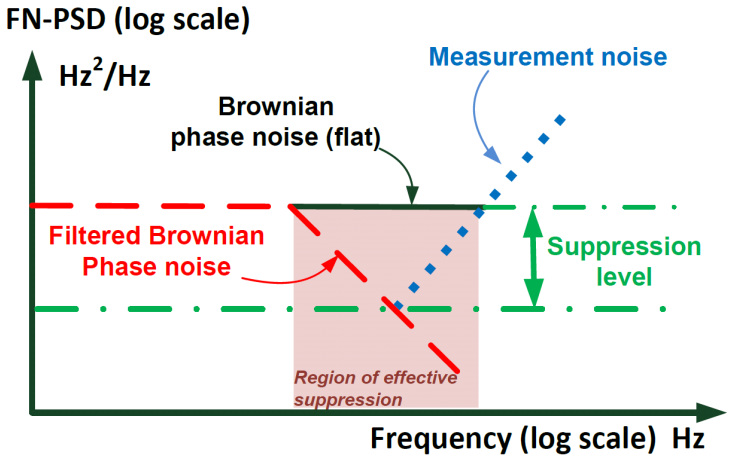

FIGURE 3.3 – Cartoon of frequency-noise power spectral density for a laser with active frequency noise reduction.

Additive white Gaussian noise (AWGN) results in a flat phase-noise power spectral density

(PN-PSD), or an upward sloping line in the FN-PSD. The AWGN FN-PSD, found in Appendix E, is

$$S_{\Delta f,AWGN}(f)_{dB} = \left[N_o \Big/ \left(2A^2\right)\right]_{dB} + 20\log_{10} f \tag{3.1}$$

where N_o is the two-sided field PSD of white Gaussian noise (flat noise floor appeared in field PSD), and A is the amplitude of electric field of the receiver signal in the absence of noise. As illustrated in Fig. 3.3 (blue dotted, upward sloping line), the f^2-FN AWGN measurement noise dominates the filtered source at high frequencies (starting from hundreds of MHz).

Following the FBG filtering, the phase noise is no longer Brownian (the frequency noise is colored), and therefore the simple parameterization by the linewidth is no longer valid, i.e., the FN-PSD is no longer flat at mid-frequencies. In particular, filtering by the ultra narrowband FBG will create a downward slope in the FN-PSD shown in Fig. 3.3 (red dash).

As the ITLA laser linewidth is inappropriate to characterize the filtered ITLA, we must turn to other parameters. One definition of the suppression accorded by the noise filtering is to measure the reduction in FN-PSD at the point where measurement noise obscures laser phase noise, as illustrated in Fig. 3.3 (green, dot-dashed lines). This parameterization, however, does not capture the frequency dependence of FN-PSD. The objective of this Chapter is to quantify the performance improvement for BER in the presence of parallelization in phase tracking. The level of parallelization will determine which frequency region will dominate. Greater parallelization level will lead to FN-PSD in the shaded region of Fig. 3.3 playing a greater role in overall performance.

TeraXion PS-TNLs provides two modes of operation :

1. a native mode, with the FBG at full transmission (without filtering) ;
2. a low-noise mode, with FBG filtering starting at 10 MHz.

The FN-PSD of the two operating modes of the PS-TNL are shown in Fig. 3.4 as measured with a self-homodyne coherent detection setup shown in Appendix F. Note that switching to the low-noise mode of PS-TNLs reduces laser phase noise, but the phase-noise statistics are changed so that the linewidth is no longer uniquely defined (refer to Appendix B). In Fig. 3.4, the upper curve shows the FN-PSD of the free-running laser without FBG filtering, and the lower curve shows that optical filtering by an FBG suppresses the white FN by 4 dB at frequencies above 50 MHz. The f^2-FN (upward sloping) noise does not come from the lasers themselves, but mainly from the electrical noise of our coherent receiver (please refer to Appendix E).

3.4 Insight from PSD

To maintain phase tracking, the phase-recovered symbols should fall well inside the decision boundary. The decision boundary for 64-QAM is defined as the threshold angle beyond which

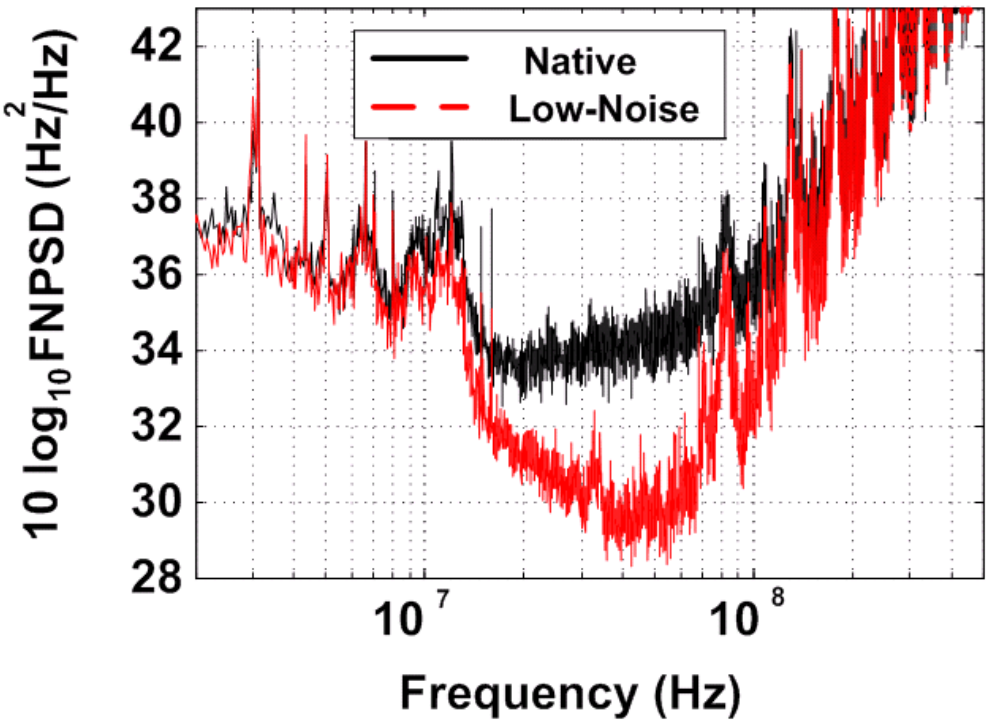

FIGURE 3.4 – FN-PSD of native mode (without FBG) and low-noise mode (with FBG).

erroneous decision will happen. This threshold angle is calculated as the distance between the corner symbol of 64-QAM constellation and its closest symbol : 4.73 degrees (0.0263π radian). As previously mentioned, the feedback loop of parallel DD-PR introduces a delay between the phase estimate and the current symbol, the phase tracking loop bandwidth for DSP in electronics becomes $R_s/(d \times P)$.

We propose a heuristic criterion to avoid loss of lock in phase tracking at moderate or high OSNR : the phase-noise increment between two adjacent symbols in each parallelization channel should be smaller than a certain threshold which is a fraction of 64-QAM decision boundary. The phase-noise increment over the response time of the feedback loop T_{loop} is proportional to the root-mean square of its variance

$$0.0263\pi/K \geq \sqrt{4\pi^2 S(f = \frac{R_s}{Pd}) T_{loop}} \tag{3.2}$$

where T_{loop} is approximately equal to $d \times P \times Tsym$, K is a proportionality constant, whose value depends on the desired BER level at high OSNR. K is usually larger than unity to make the threshold tighter in the presence of AWGN phase noise. Note that for phase tracking, we should observe FN-PSD at operating frequency (loop bandwidth) $f = R_s/(d \times P)$, which is

the corner frequency beyond which the feedback loop starts to lose tracking. (3.2) becomes

$$S_{\Delta f}(f = \frac{R_s}{Pd}) \leq \frac{(0.0263)^2/K}{4PdT_{sym}} \tag{3.3}$$

Equation (3.3) tells us that when the parallelization level increases, the loop bandwidth of parallel DD-MLE decreases, and the FN-PSD level has to be smaller by a factor of P in order to maintain phase tracking. For filtered sources, the FN-PSD is shaped by the FBG filter with a corner frequency smaller than the phase-tracking loop bandwidth. The FN-PSD level, due to the FBG suppression, is lower than that of the original laser within the loop bandwidth (as shown in the red-shaped region in Fig. 3.3, or 10-100 MHz in Fig. 3.4). Therefore filtered lasers can allow higher parallelization for a fixed number of pipelining delay elements on the feedback path.

3.5 System Performance

3.5.1 Experimental Setup

In this section, we compare the bit-error rate (BER) performance of FBG-filtered lasers (low-noise mode) with that of unfiltered lasers (native mode) using parallel phase tracking in a time-interleaving structure as discussed in Section 2.4. Fig. 3.5 shows the experimental setup for single polarization back-to-back 5 Gbaud 64-QAM. Two separate PS-TNLs tuned at 1550 nm were used as transmitter and local oscillator sources. To examine the effect of FBG filtering, two different measurements were taken :

1. both sources in native mode ;
2. both sources in low-noise mode.

A 20-GSa/s, 6-bit arbitrary waveform generator (AWG) operating at 4 samples per symbol was used to generate two 5 Gbaud 8-level electrical signals. Data was a repeated sequence of 98304 bits (limited by AWG memory) taken from a pseudorandom bit sequence with a word length of 31, driving the in-phase/quadrature (IQ) modulator. Wiener-Hopf-based [48] predistortion was applied to compensate for RF components and the imperfect linear gain of the power amplifier in the transmitter.

A combination of variable optical attenuator and EDFA was used to adjust the received OSNR. To maintain the integrated coherent receiver at its optimum operating point, the received optical power was fixed at -8 dBm and the local oscillator power was set to 13.8 dBm. The coherently detected signal was sampled by the real time oscilloscope at 80 Gsa/s with 30-GHz electrical bandwidth. The captured samples were then retimed and resampled to one sample per symbol for the subsequent DSP.

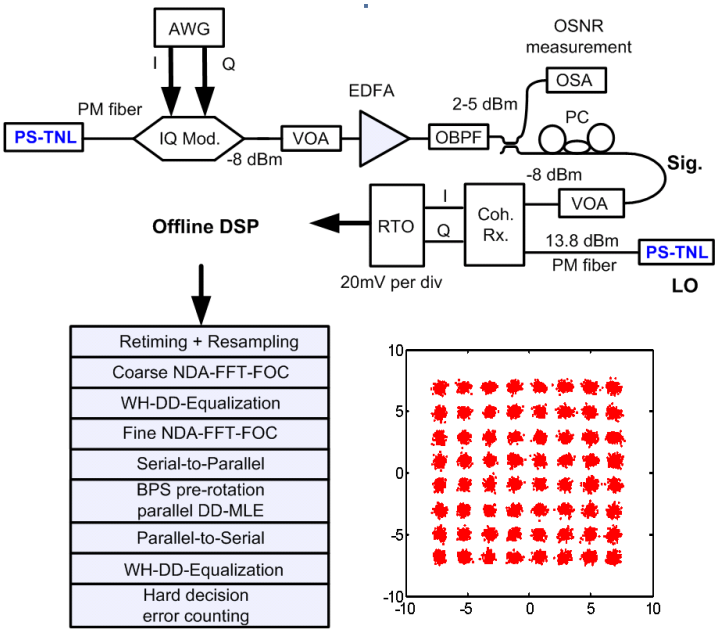

FIGURE 3.5 – Top : experimental setup of back-to-back 5-Gbaud 64-QAM. Right bottom : Recovered 64 QAM constellation without parallelization at OSNR = 28 dB (BER= 6e-5). IQ mod : In-phase/Quadrature modulator, AWG : Arbitrary waveform generator, PM fiber : Polarization maintaining fiber, VOA : variable optical attenuator, OBPF : optical bandpass filter, PC : polarization controller. Coh. Rx. : Coherent Receiver. RTO : Real time oscilloscope. EDFA : Erbium doped fiber amplifier.

3.5.2 Offline DSP

A non-data-aided fast Fourier transform-based frequency offset compensation (NDA-FFT-FOC) [53] was performed in a block-wise fashion over 30,000 symbols, to coarsely remove large frequency offset (in GHz) between the two lasers at transmitter and at receiver. We applied training-sequence-based Wiener-Hopf-based decision-directed equalizer (WH-DD-EQ) with 31 taps [48]. Fine NDA-FFT-FOC was performed to further remove the effect of frequency offset dynamics [69]. The serial data was demultiplexed into P channels, where P was varied from 8 to 30 to observe the performance of different levels of parallelization (recall that the total number of symbol delays due to the feedback loop is approximately $P \times d$).

Blind phase search (BPS)-based pre-rotation [69] was used to obtain the initial phase of the first symbol of each channel. Subsequently, parallel DD-MLE (see Section 2.4) was used for phase recovery, incurring a pipelining delay of four [75]. The P channels were then recombined into a single data stream. We again applied 31-tap WH-DD-EQ [48] to further equalize with the more reliable decisions following phase recovery. Finally, hard decision was performed on I and Q individually, and we counted errors. BER was estimated over 7 114 200 bits for each OSNR value, which allows reliable estimation down to BER = 5.6e-5.

3.5.3 Results

In Fig. 3.6 we present BER versus OSNR results for native mode (square markers) and low-noise mode (circle markers). For serial data processing (solid curves), the two sources offer similar performance. For parallel DD-PR with P parallelization level and d pipelining registers on feedback path, the processing rate of each channel becomes R_s/P, leading to larger phase excursions, where R_s is the optical baud rate. The feedback delay due to the d pipelining registers further reduces the loop bandwidth to approximately $R_s/P/d$, as discussed in Section 2.4. Hence, in Fig. 3.6, parallel processing leads to an increase in BER, but the FEC threshold (defined at BER = 1e-3) is still respected. For 12 parallel rails (dashed curves) and 24 rails (dotted curves) we see a reduced BER floor for the low noise mode.

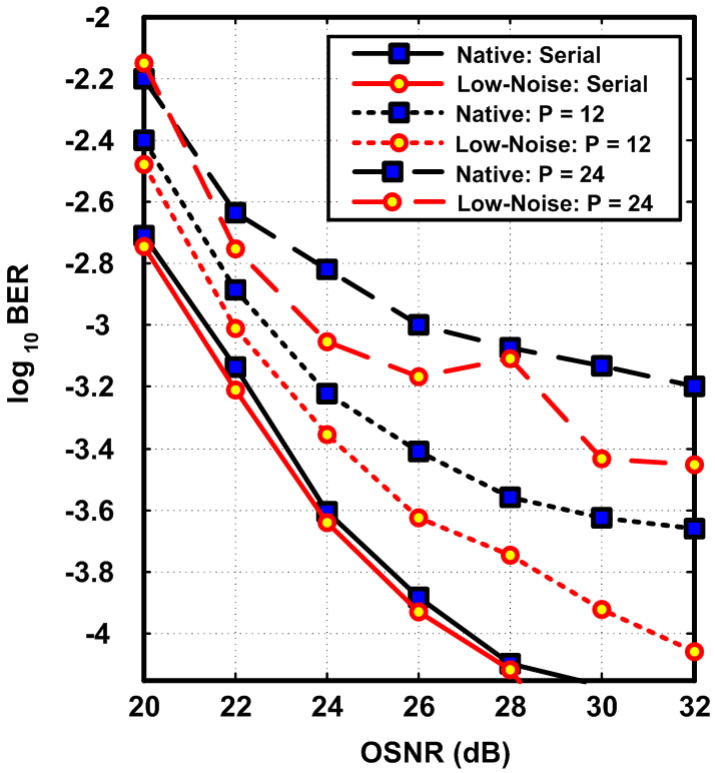

FIGURE 3.6 – BER versus OSNR for native and low-noise mode of PS-TNLs for serial and parallel phase tracking ($d = 4$) with $P = 12$ and $P = 24$.

With FBG noise suppression, the phase tracking error decreases, improving BER. Note that the value for the curve $P = 24$ at OSNR = 28 dB is unexpectedly high. It is due to the frequency instability of the TeraXion laser source [3] We swept parallelization levels P, determined

3. The TeraXion source requries an error signal to lock the CW center frequency within the passband of the FBG filter. This error signal was sinusoidal originally at 75 kHz, and this frequency exceeds the bandwidth of the parallel DD-PR as explained in Chapter 4, and therefore extra effort in tracking and compensating the

BER and calculated the OSNR penalty at BER = 1e-3 for each level as shown in Fig. 3.7. A second axis is plotted across the top of Fig. 3.7 indicating the bandwidth of the phase tracking loop for the given parallelization level. The shaded area indicates where the phase tracking bandwidth falls within the noise-suppression bandwidth of the low-noise mode (with FBG noise suppression). This area is determined by Fig. 3.4 where the low-noise-mode FN-PSD falls below the native-mode FN-PSD. We see clearly that the low-noise mode improves the BER compared to the native mode within the shaded area.

For $P = 8$ and 12, corresponding to phase tracking loop bandwidth of 156.25 MHz and 104.17 MHz, respectively, the low-noise mode cannot provide significant improvement. In this frequency region, the FN-PSD is dominated by the f^2 frequency noise (i.e., white phase noise, shown in Appendix E). As shown in Fig. 3.4, the f^2 frequency noise is visible in the linearly increasing logarithmic FN-PSD above 70 MHz. For higher parallelization, the phase tracking effective loop bandwidth decreases, entering a band where the f^2 frequency noise level of the coherent receiver is lower than the white frequency noise level of the native mode (unfiltered laser)

The frequency noise of laser sources between 10 MHz and 100 MHz affects the real-time decision-directed phase recovery for 64-QAM at the symbol rate (5 Gbaud) examined. At this symbol rate, the tracking bandwidth due to hardware parallelization and pipelining falls within the noise suppression band for a wide range of parallelization levels (above 12). This conclusion should hold for higher symbol rates as well, as higher symbol rates require higher parallelization level, reducing the tracking bandwidth. The low-noise mode shows a 2 dB improvement over the native mode for $P > 24$. When using the native mode, parallel DD-PR even fails for $P > 26$ while FBG filtering allows the break down point to extend up to $P = 30$.

Fig. 3.8(a) and Fig. 3.8(b) shows the recovered constellation diagrams of native mode and low-noise mode, respectively, for $P = 26$ at OSNR = 32 dB. Fig. 3.8(a) shows the rotation of symbols at the outermost ring, leading to high bit error rate (BER = 1e-3). With FBG suppression, the FN-PSD is lowered, helping the reduction of rotation of symbols at the outermost ring (BER = 4e-4) as shown in Fig. 3.8(b). As mentioned in Chapter 2, power consumption of CMOS chips increases linearly with clock frequency [42], we compare $P = 16$ for native mode with $P = 20$ for low-noise mode for the same OSNR penalty in Fig. 3.7. The decrease of processing rate from 312.5 MHz to 250 MHz allows a 6.3 % reduction in power consumption for performing phase recovery in hardware (Appendix C).

frequency offset dynamics was performed before the parallel DD-PR, which was not shown in this thesis. The unexpectedly large BER value appeared on the curve $P = 24$ at OSNR = 28 dB was caused by the failure of compensating an occasionally extreme frequency offset dynamics (even higher than 75 kHz) caused by the frequency control of the laser.

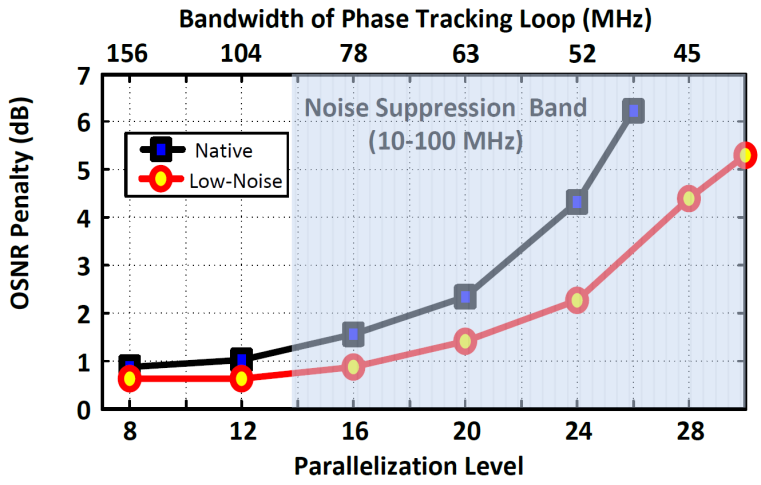

FIGURE 3.7 – OSNR penalty for back-to-back 5 Gbaud 64-QAM with $d = 4$ versus parallelization (lower axis) and versus bandwidth of phase tracking loop (upper axis) for both native mode and low-noise mode. Shaded area is the noise suppression band, i.e., in Fig. 3.4 where low-noise FN-PSD falls below native FN-PSD.

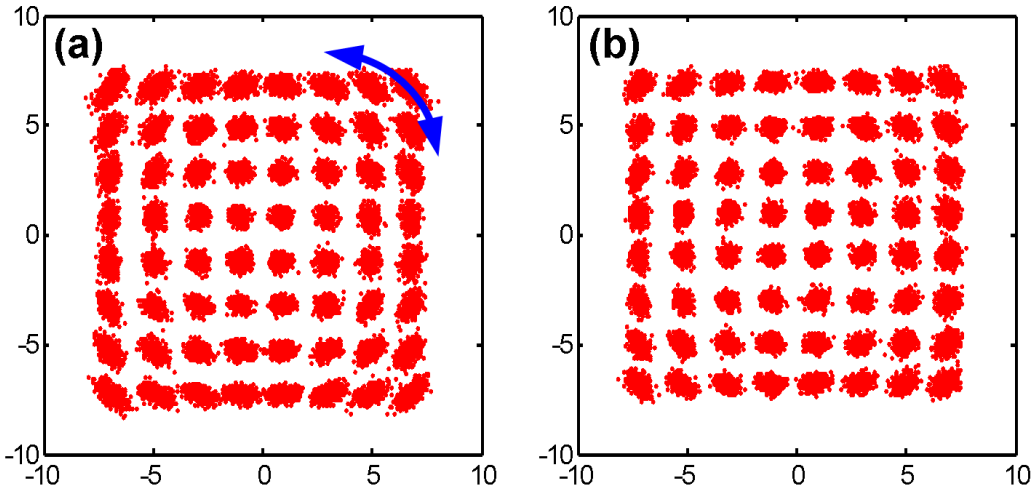

FIGURE 3.8 – Constellations over 30 000 symbols, $P = 26$, $d = 4$, OSNR = 32 dB for (a) native mode (BER = 1e-3) and (b) low-noise mode (BER = 4e-4).

3.6 Conclusions

In this chapter, we have seen that laser linewidth is no longer appropriate to quantify laser phase noise. Frequency noise power spectral density levels are better descriptors of the impact of laser phase noise for real-time phase tracking with feedback delay, in contrary to nowadays

offline experimental demonstrations in which the absence of feedback delay is assumed. We experimentally showed the system performance given by TeraXion's narrow-linewidth lasers on parallel and pipelined DD-MLE in a single-carrier 5 Gbaud 64-QAM back-to-back heterodyne coherent system. Using a FBG filter (an optical approach) further suppresses the frequency noise spectral level of contemporary narrow-linewidth semiconductor lasers such as ECL or ITLA. We found that the frequency noise of laser sources between 10 MHz and 100 MHz affects the real-time decision-directed phase recovery. For parallelization level higher than 24, TeraXion's narrow-linewidth lasers show more than 2 dB improvement in OSNR penalty compared to that of conventional lasers. For the same OSNR penalty, the optically-filtered laser permits greater parallelization, e.g., increase from 16 to 20, to reduce the hardware processing rate from 312.5 to 250 MHz.

Chapitre 4

Impact of Sinusoidal Phase Modulation on Phase Tracking

In Chapter 2, we found that real-time implementation of DSP at Gbaud rates requires hardware parallelization to increase the effective sampling rate of each sub-data rate parallel channel and to reduce power consumption [42]. In Chapter 3, we have seen that the frequency noise of laser sources between 10 MHz and 100 MHz affects the real-time decision-directed phase recovery as a result of reduction of tracking bandwidth due to hardware parallelization and pipelining. In Chapter 3 we assumed no frequency offset (or after a perfect frequency offset compensation). For systems employing dense constellations, such as 64-QAM, the phase margin (the allowable maximum amount of phase rotation without tracking loss) is greatly reduced compared to the commercially available 16-QAM. A residual frequency offset will give a phase rotation larger than the pure Wiener phase noise of laser sources. Because of the bandwidth-reduced phase tracking, we may foresee difficulties in implementing 64-QAM coherent receivers. Nevertheless, the impact of the non-ideal frequency offset compensation on the real-time phase tracking with feedback loop has not been investigated. This chapter is devoted to explore the phase tracking failure caused by the frequency components of laser sources below 10 MHz.

We will revisit a best-known, standard frequency offset compensation algorithm, followed by a detailed explanation of the impact of residual frequency offset on real-time phase tracking. We will present our experimental results on the effect of sinusoidal laser phase noise on parallel decision-directed maximum likelihood estimation (DD-MLE) with four pipelining delays on its feedback path [75] in a single-polarization [1] single-carrier 5 Gbaud 64-QAM system. We use simulation to parameterize the range of frequency modulation-amplitude and frequency-

1. We only used a single-polarization experimental setup because the dual-polarization setup was not ready by the time of the publication related to this chapter. Moreover, this chapter is only devoted to the impact of sinusoidal phase modulation on phase tracking, and therefore this single-polarization back-to-back experiment isolated the phase noise effect from polarization impairment.

modulation frequency that lead to tracking failure. Taking into account the effects of the conventional frequency offset estimation algorithm as well as equalization, we examine phase tracking performance as a function of parallelization level.

4.1 Introduction

Frequency offset between transmit and receive lasers can be interpreted as a long-term linear phase change. If not correctly compensated this offset can disrupt phase tracking. Non-data-aided fast-Fourier-transform (NDA-FFT)-based, also called periodogram-based, frequency offset compensation (FOC) [53] is typically preferred over the less practical data-aided phase-locked loop (PLL) approach [49, 76]. NDA-FFT gives one single frequency-offset estimate [2] over the data block of concern, and its feedforward structure facilitates hardware implementation. The PLL allows tracking [3] of frequency-offset dynamics but is impractical due to large feedback delay (which requires the feedback of phase estimates to the FOC stage) [49, 76], and thus can only be used for offline demonstrations, but not for real-time implementations. The state-of-the-art FOC algorithms for QPSK and 16-QAM were designed to compensate a large frequency offset. The small residual (or uncompensated) frequency offset is assumed to be canceled in the subsequent feedforward phase recovery like (modified) Viterbi-Viterbi algorithm in Section 2.2.2 or blind phase search in Section 2.3.1. However, the small residual frequency offset may exceed the phase margin of phase tracking for denser constellations such as 64-QAM.

For offline serial phase tracking [43, 76, 18], residual frequency offset or fast frequency modulation (FM) does not affect decision-directed phase tracking (DD-PR) as the bandwidth of the tracking loop exceeds that of phase variation. For real-time implementation, parallelization and pipelining introduce a large delay on the feedback path of DD-PR, reducing the bandwidth of parallel phase tracking as mentioned in Chapter 2. Residual frequency offset or fast FM creates a phase rotation that causes tracking failure for parallel phase recovery. In the presence of sinusoidal FM, the varying phase spoofs the NDA-FFT-based FOC, leading to sporadic phase tracking failure. Data blocks falling in certain sections of the sinusoidal cycle with fast phase variation, i.e. near sinusoidal maxima, are prone to tracking failure (to be explained later).

Sinusoidal frequency modulation may be intentional for control purposes, or incidental due to electronics and environmental fluctuations. For instance, parasitic sine tones can originate in switching power supplies and power converters driving the laser diode. Kuschnerov et al. [30] experimentally investigated the impact of mechanical vibrations using sinusoidal FM in a laser source. That study examined QPSK with Viterbi-Viterbi (V-V) phase tracking algorithm

2. NDA-FFT belongs to parameter estimation. Please refer to Appendix A.
3. The PLL for frequency offset tracking does not require the temporal dynamics of the frequency offset. Please refer to Appendix A.

as well as 16-QAM with blind phase search. They showed that the FM amplitude and FM frequency should be smaller than 120 MHz and 35 MHz, respectively, for QPSK and 16-QAM for penalty to be negligible.

Gianni et al. [17] analyzed 10 GBaud QPSK via simulation a parallel second-order DPLL (digital phase locked loop) followed by V-V algorithm in the presence of frequency modulation, extending [16] to 16-QAM by adding one more BPS stage. Qiu et al [49] recently proposed using shorter data blocks for QPSK and 16-QAM for frequency offset tracking to compensate rapid drift. Their approach is only applicable to offline processing, as they did not take into account feedback delay in parallel DPLL. There has been no report to date, to the best of our knowledge, of the impact of FM in 64-QAM, where phase-error tolerance is greatly reduced vis-à-vis 16-QAM.

Contribution

We will investigate the impact of the sine tones of laser sources on parallel and pipelined decision-directed phase recovery (maximum likelihood estimation) in digital coherent receivers, as a result of residual frequency offset uncompensated by the conventional frequency offset compensation algorithm (NDA-FFT)[53]. Experimentally, we limit our investigation to a back-to-back, single-polarization, single-carrier 5 Gbaud 64-QAM heterodyne coherent system as shown in Fig. 4.1. We have made the following contributions in this chapter :

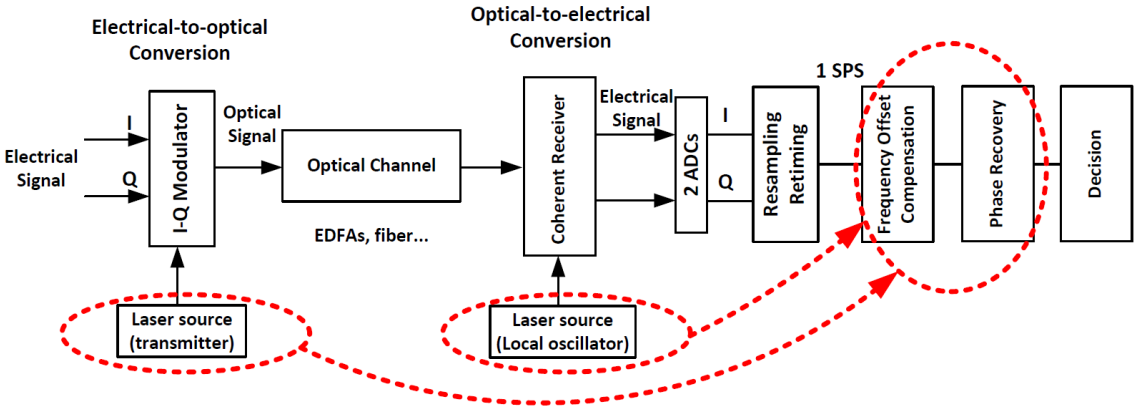

FIGURE 4.1 – Block diagram of a single-polarization, single-carrier coherent system.

1. We pointed out, for the first time, that the standard NDA-FFT FOC algorithm could be the cause of real-time phase tracking failure for 64-QAM in the presence of sine tones in laser sources, whose frequencies lie on the range below 1 MHz, shown in Fig. 2.8.
2. We investigated, for the first time, the ranges of the FM amplitude and FM frequency of laser sources to avoid real-time phase tracking failure using simulation.

4.2 Frequency Offset Compensation Algorithm

In this section, we will briefly introduce the frequency offset compensation algorithm that we will consider in this chapter.

Selmi et al., first applied the non-data-aided fast-Fourier-transform (NDA-FFT) method, also called periodogram-based or FFT-maximization-based method, for frequency offset compensation (FOC) [53]. The frequency offset $\Delta\hat{f}_{FO}$ is estimated by

$$\Delta\hat{f}_{FO} = \frac{1}{4}\max_{f}\left\{\text{fft}(r^4[k])\right\}, \tag{4.1}$$

where fft denotes the fast Fourier transform over a certain block size L_{FFT}. $r[k]$ is the received symbol, where k is the time index taking over L_{FFT}. Therefore, frequency offset estimate $\Delta\hat{f}_{FO}$ is assumed to be constant over every L_{FFT} symbols. This frequency offset estimation becomes inaccurate when the frequency offset dynamics is fast within the block. Practically, to reduce the computational power, k is taken for every symbol duration T_{sym} rather than the sample duration because the sampling rate of a system is usually higher than the symbol rate.

NDA-FFT FOC raises the received signal (which includes a true frequency offset Δf_{FO}) by a power of four to remove the symmetry of QAM signals (to avoid ambiguities in frequency offset estimation), and then to find out the frequency component, f, such that the phasor $\exp(j2\pi fkT_{sym})$ results in a maximum correlation with the 4^{th}-power received signal, $r^4[k]$, over a block length (a time duration). This is equivalent to a FFT operation. In the FFT spectrum, the maximum peak refers to a frequency value close to $4\times\Delta f_{LO}$. This value is then divided by 4 to obtain the frequency offset estimate, $\Delta\hat{f}_{FO}$. Then we multiply the received signal $r[k]$ symbol-by-symbol by $\exp(-j2\pi\Delta\hat{f}_{FO}kT_{sym})$.

4.3 Effect of Residual frequency offset

After having discussed the NDA-FFT FOC algorithm, this section will explain in detail the impact of residual frequency offset on real-time phase recovery with feedback with the aid of our simulation. First of all, we parameterize the sine tones in our laser sources (in both transmitter and receiver) with sinusoidal FM modulation at frequency f_m and peak-to-peak FM amplitude of A_{pp}. The electric field $A(t)$ can be expressed as

$$A\left(t\right) = \exp\left\{j2\pi f_c t + j\frac{A_{pp}}{2f_m}\cos\left(2\pi f_m t\right) + j\theta_{LPN}\left(t\right)\right\} \tag{4.2}$$

where f_c is the laser carrier frequency and θ_{LPN} is the laser phase noise.

In order to explain the effect of residual frequency offset on the parallel and pipelined maximum likelihood phase estimation, only the temporal change of laser phase noise, excluding

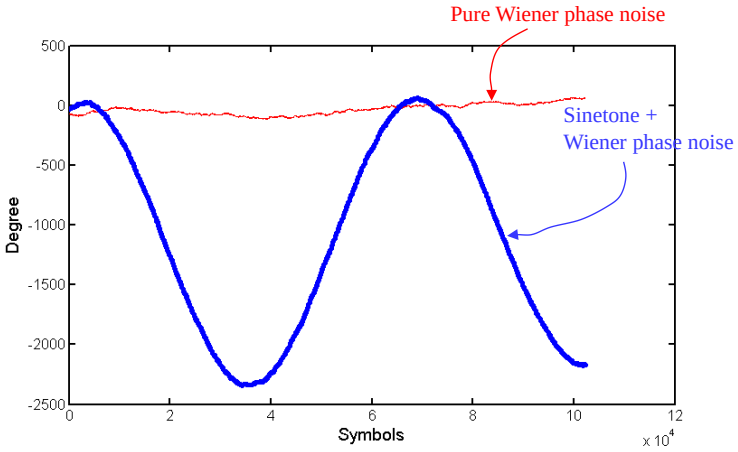

FIGURE 4.2 – Simulated laser phase. Blue : sine tone with Wiener phase noise ; Red : pure Wiener phase noise

the data modulation, is used to pass through the FFT-FOC. Let us choose $f_m = 75$ kHz, $A_{pp} = 1.5$ MHz. Fig. 4.2 shows the phase evolution of a sine tone together with the laser phase noise for a total linewidth of 20 kHz generated using the method shown in Appendix B. The phase is obtained by taking the argument of (4.2). It is obvious that the peak-to-peak amplitude of the sine tone dominates the Wiener phase noise.

FFT-FOC is performed over L_{DS} symbol, where L_{DS} is the length of FFT block size used to undergo FOC. We separate into the following three cases for illustrating the impact of sine tones :

1. $L_{DS} = 8000$ **taken away from the sine tone's maxima** : In Fig. 4.3a, the FFT block covers the entire region of a linear phase [4] equal to $2\pi\Delta f_{FO} n T_{sym}$, where n is the time index and T_{sym} is the symbol duration. The frequency offset estimation is based on the maximization of periodogram of the 4^{th}-power received QAM signal, which gives the frequency offset value, Δf_{FO}, resulting in a highest correlation with the signal content. In Fig. 4.3b, after FOC, the covered section is de-rotated by $e^{-j2\pi\Delta f_{FO} n T_{sym}}$. The large phase change is thus compensated, while the small-varying phase fluctuation is due to Wiener phase noise.

2. $L_{DS} = 8000$ **taken close to the sine tone's maxima** : In Fig. 4.4a, the FFT block covers partially the sine tone maxima. FFT-FOC gives a single estimated value based on where the FFT block covers, Δf_{FO}, corresponding to the phase change only in the region showing linear phase change. In Fig. 4.4b, after FOC, regions of linear phase change and sine tone maximum are both derotated by $e^{-j2\pi\Delta f_{FO} n T_{sym}}$. The region showing linear

4. "Linear phase" refers to the phase changing with time linearly, i.e., with a constant slope equal to $2\pi\Delta f_{FO}$.

phase change is compensated and only small-varying phase change due to Wiener phase noise remains. However, the region covering the sine tone maximum is over-rotated, resulting in a larger phase change.

3. $L_{DS} = 12000$ **taken close to the sine tone's maxima** : In Fig. 4.5a, the FFT block covers partially the sine tone maxima, but it does more than the case of $L_{DS} = 8000$. FFT-FOC gives a single estimated value, Δf_{FO}, corresponding to the phase change only in the region showing linear phase change. In Fig. 4.5b, after FOC, the region covering the sine tone maximum is over-rotated, resulting in a very large phase change that exceeds the phase margin for parallel phase tracking described by (3.3).

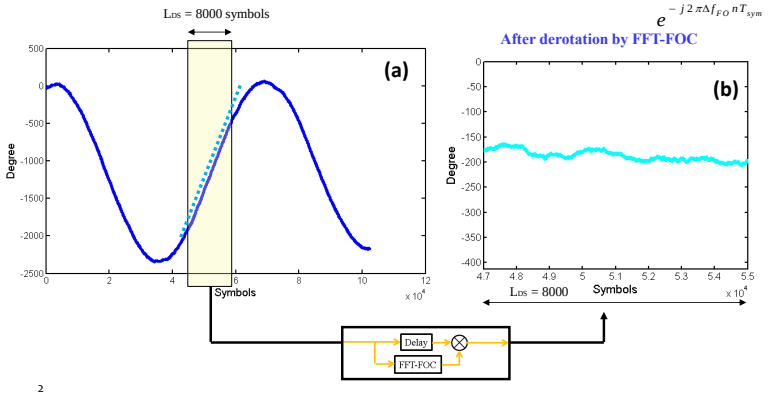

FIGURE 4.3 – Simulated impact of FOC within a time window of $L_{DS} = 8000$ symbols chosen in the linear region of the laser phase. Left : before FOC ; Right : After FOC.

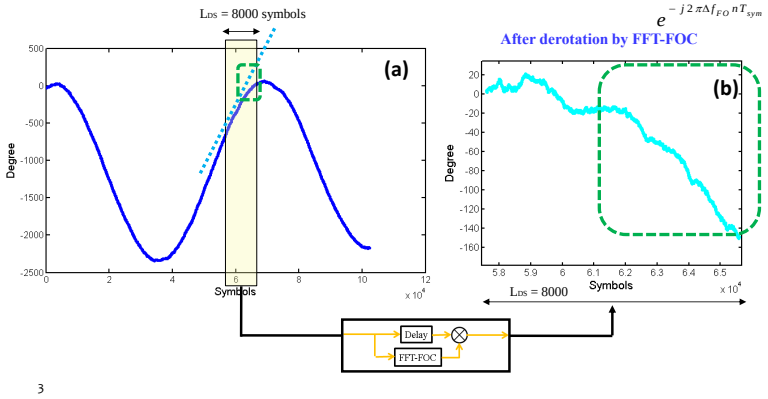

FIGURE 4.4 – Simulated impact of FOC within a time window of $L_{DS} = 8000$ symbols chosen in proximity of the sine tone maxima. Left : before FOC ; Right : After FOC.

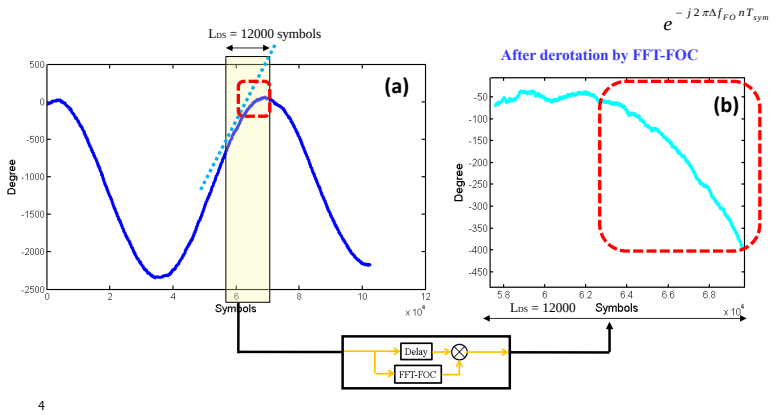

FIGURE 4.5 – Simulated impact of FOC within a time window of $L_{DS} = 12000$ symbols chosen in proximity of the sine tone maxima. Left : before FOC ; Right : After FOC.

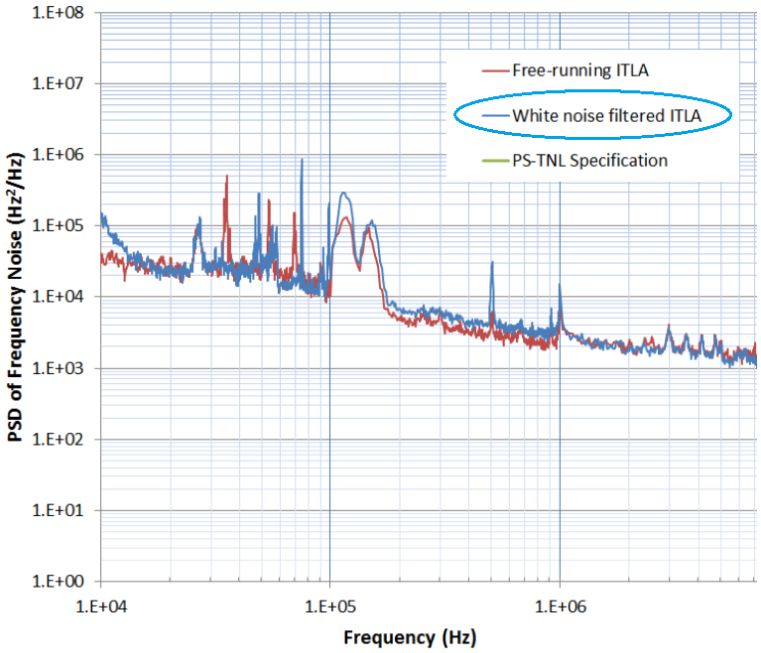

FIGURE 4.6 – Frequency noise PSD of our laser under test (blue : white noise filtered ITLA). Source : TeraXion PS-TNL specification.

4.4 System Perfomance

4.4.1 Experimental Setup

The source provided by TeraXion had three settings

1. no sinusoidal modulation,
2. a sinusoidal FM with $f_m = 25$ kHz, $A_{pp} = 0.6$ MHz,
3. a sinusoidal FM with $f_m = 75$ kHz, $A_{pp} = 1.5$ MHz.

Due to device limitations [5] , we were unable to tune f_m and A_{pp}, thus we examined the three cases permitted, on both (transmit and local oscillator) 10 kHz linewidth laser sources.

We performed an experiment for back-to-back 5 Gbaud 64-QAM with the setup used in Fig. 4.7. A 20-GSa/s, 6-bit arbitrary waveform generator (AWG) operating at 4 samples per symbol was used to generate two 5-Gbaud 8-level electrical signals. Data was a repeated sequence of 98304 bits (limited by AWG memory) taken from a pseudo-random bit sequence of length 2^{31}-1, driving the in-phase (I) quadrature-phase (Q) modulator. Wiener-Hopf-based pre-distortion was applied to compensate for limited bandwidth RF components as well as nonlinear gain of the power amplifier at the transmitter. A combination of a variable optical attenuator and an EDFA was used to adjust the received OSNR.

The coherently detected signal was sampled by a real time oscilloscope with 30 GHz electrical bandwidth sampling at 80 Gsa/s. The captured samples were retimed and resampled to one sample per symbol.

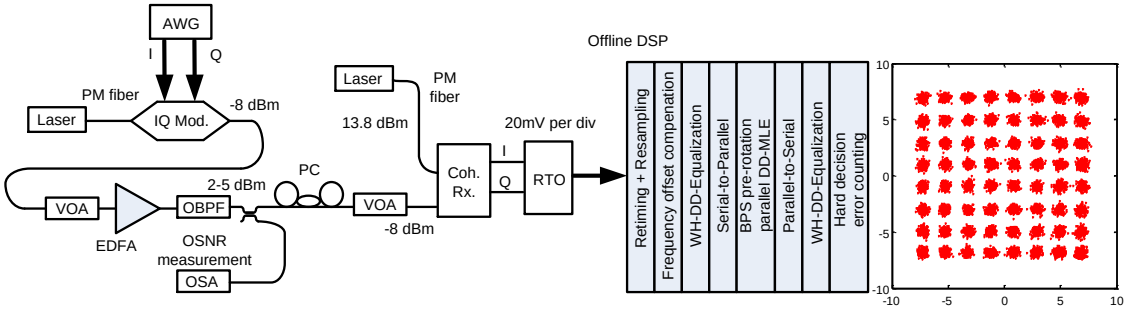

FIGURE 4.7 – Experimental setup of back-to-back 5-Gbaud 64-QAM. IQ Mod. : in phase-quadrature modulator, AWG : Arbitrary waveform generator, PM : polarization maintaining, VOA : variable optical attenuator, OBPF : optical bandpass filter. PC : polarization controller. Coh. Rx. : Coherent Receiver. RTO : Real time oscilloscope. EDFA : Erbium doped fiber amplifier. Right : Recovered 64 QAM constellation without parallelization at OSNR = 28 dB (BER = 6e-5).

5. Following instructions provided by TeraXion, we tried to manually tune fm and App to other settings than those pre-programmed. Unfortunately, these two parameters are interdependent, and they vary with temperature as well (manual tuning mode switches off temperature control). We were therefore only able to run our experiment using the three settings : no sine tone, sine tone at $f_m = 25$ kHz, $A_{pp} = 0.6$ MHz, and sine tone at $f_m = 75$ kHz, $A_{pp} = 1.5$ MHz.

4.4.2 Offline DSP

To investigate the impact of block length on FO estimation, NDA-FFT-FOC [53] with zero padding (to guarantee a fixed FFT resolution) was performed repeatedly over data segments of L_{DS} symbols, where L_{DS} (length of data segment) = 4k, 6k, 8k 10k and 12k. For QPSK and 16-QAM, the transmitter and receiver limitations are not too stringent and equalization may not be required [18, 4] . However, for 64-QAM, equalization becomes necessary to achieve a clear constellation and good bit-error rate (BER). We employed the statically optimal equalization (compared to adaptive equalization [76, 18]), in the form of a Wiener-Hopf-based decision-directed equalizer (WH-DD-EQ) with 31 taps (refer to Section 3.5.2). The WH-DD-EQ was updated by the same data segment of L_{DS} symbols at the output of NDA-FFT- FOC.

The serial data was demultiplexed into P channels, where P was varied among 4, 6, 8, 10 and 12 to observe the performance of different levels of parallelization. 64-QAM is much more sensitive to initial phase offset estimation than that for 16-QAM or QPSK when using DD-phase tracking. To solve this problem, BPS-based pre-rotation (in Section 2.3.1) was used to obtain a good initial phase offset estimate at the first symbol in each parallel channel. Subsequently, parallel DD-MLE incurring four pipelining delays was used for phase recovery. The total delay is $4P$, the product of pipelining delays and parallelization. The P channels were then recombined into a single data stream. We again applied 31-tap WH-DD-EQ to further equalize with the more reliable decisions following phase recovery. Finally, hard decisions were performed for I and Q individually, and we counted errors. At each OSNR, 7114200 bits were examined, allowing reliable BER estimation down to 5.6e-5.

4.4.3 Results

The results of BER versus OSNR of parallel DD-MLE were compared with that of serial DD-MLE to obtain the OSNR penalty at BER = 1e-3. Fig. 4.8 shows the experimental OSNR penalty versus different levels of parallelization, P. Each curve corresponds to a different length of data segment L_{DS} for performing NDA-FFT-FOC and WH-DD-EQ. From Fig. 4.8, we can understand :

- **Fixed P and smaller L_{DS}** : smaller L_{DS} for NDA-FFT-FOC to reduce the fast varying phase change due to the sine tone, but leads to an inadequate number of symbols for training the equalizer.
- **Fixed P and larger L_{DS}** : a larger L_{DS} results in a lower OSNR penalty as more symbols can be used to train the WH-based equalizer for better channel estimation. Equalization does not show further improvement for $L_{DS} > 12k$.
- **Fixed L_{DS} and smaller P** : For P smaller than 8, OSNR penalty increases as P decreases, because smaller P leads to insufficient moving average in the parallel imple-

mentation of DD-MLE, leading to poor phase tracking (the number of symbols in the moving average is proportional to P) [75].

– **Fixed L_{DS} and larger P** : Tracking failure was observed for $P = 12$ or higher. We attribute this to the spurious small-amplitude fast frequency components between 100 kHz and 300 kHz. The origin of these spurious tones is unknown and they do not appear in the specification of the laser sources under test.

From the above, we can see that **in the presence of fast changing frequency offset or FM, it is better to keep the block length for NDA-FFT-FOC short, but long enough to give good equalization**. For the performance comparison between three cases, we can observe :

– **Optima parallellization level** : The optimal P is 8, corresponding to a processing rate of 625 MHz.

– **Case 1 and Case 2** : Fig. 4.8(b) shows that, in the presence of sinusoidal FM with $f_m = 25$ kHz, $A_{pp} = 0.6$ MHz, the performance is similar to that without FM shown in Fig. 4.8(a).

– **Case 3** : In the presence of sinusoidal FM with $f_m = 75$ kHz and $A_{pp} = 1.5$ MHz, however, loss of tracking was observed for all values of P, resulting in high OSNR penalty, as shown in Fig. Fig. 4.8(c). This can be explained by the behavior in the following subsection, documented in Fig. 4.9.

Experimental Proof for Tracking Failure

Fig. 4.9(a) shows the phase variation of a sinusoidal FM with $f_m = 75$ kHz, $A_{pp} = 1.5$ MHz added to a 10 kHz-linewidth laser source. The laser phase noise is tracked over 12 μs (a duration of 60000 symbols at 5 Gbaud). The true phase (in blue in plots in Fig. 4.9) was obtained through knowledge of transmitted symbols. The data segment includes a fast varying phase change as shown in Fig. 4.9(b), corresponding to the shaded area in Fig. 4.9(a). The NDA-FFT-FOC contains a spectrum (periodogram) that mainly contains two frequency components, namely, the estimated frequency offset, $\Delta f = 922$ kHz from the first half of Fig. 4.9(b) corresponding to the positive linear slope (332 degrees per μs), and the DC components from the second half of Fig. 4.9(b) corresponding to the constant-phase region. Since the frequency component of $\Delta f = 922$ kHz dominates the DC component, the maximization of the periodogram (the principle of the NDA-FFT-FOC) outputs the estimated frequency offset, $\Delta f = 922$ kHz, and derotates the whole data segment by $2\pi\Delta f t$.

The resultant FO-compensated data segment is shown in Fig. 4.9(c), in which the first half remains flat while the second half is over-compensated and has a linear slope of .316 degrees per is. The serial DD-PR works well (in red), while the parallel DD-phase tracking algorithm cannot keep up with the fast decreasing phase change, and therefore tracking fails (in green). One may suggest the usage of fine FOC, i.e. cascading one more FOC stage with short block

length to cancel the sine tone, but it causes additional power consumption in implementing real-time products.

4.4.4 Simulation Results

In this section, we are going to discuss our simulation results for sweeping FM frequency f_m and FM amplitude A_{pp} to see OSNR penalties, which allows engineers to know how we can avoid phase tracking problems in 64-QAM real-time systems.

Since our source did not permit tweaking of f_m or A_{pp}, we swept the FM frequency f_m from 0 kHz to 95 kHz, the FM amplitude A_{pp} from 0 Hz to 2 MHz. The data segment L_{DS} for frequency offset estimation and equalization was fixed at 8000 as suggested by the experimental results that this number will give sufficient performance for all cases. The parallelization level P was taken at 8, 10, 12 (we are not interested in $P = 4$ and 6 as low parallelization level requires faster ASIC or FPGA). The results are shown in Fig. 4.10.

Fig. 4.10a ($P = 8$) shows that our experimental case 2 ($f_m = 25$ kHz, $A_{pp} = 0.6$ MHz) results around 2 dB penalty, which makes no difference with that without sine tone ($f_m = 0$ kHz, $A_{pp} = 0$ MHz), while our experimental case 3 ($f_m = 75$ kHz, $A_{pp} = 1.5$ MHz) results in more than 3 dB penalty.

Fig. 4.10b ($P = 12$) shows the our experimental cases 1 and 2 still performs similar to $P = 8$, while our experimental case 3 results in more than 4.5 dB penalty (tracking failure always happens).

Fig. 4.10b ($P = 12$) shows that our experimental case 2 results 2.5-3 dB penalty while our experimental case 3 shows tracking failure.

We see that our numerical simulation results agree with our experimental results. To avoid the extra penalty due to tracking failure, our numerical results suggest that the corresponding combinations of f_m and A_{pp} should be chosen inside the dark region of the lowest penalty in Fig. 4.10a. While simulation results are optimistic, they provide guidance on sensitivity to sine tone parameters. In particular, the simulations indicate parameter sets leading to tracking failure.

4.5 Conclusions

In this chapter, we first explained the concept the standard NDA-FFT FOC algorithm. Then we have illustrated that the standard NDA-FFT FOC algorithm could be the cause of real-time phase tracking failure for 64-QAM in the presence of sine tones in laser sources, whose frequencies lie on the range below 1 MHz. We demonstrated experimentally the impact of sinusoidal laser phase noise in the presence of parallelization and pipelining delay. Our de-

monstration applied parallel and pipelined DD-MLE to a 5-Gbaud 64-QAM system, taking into account the popular NDA-FFT for frequency-offset compensation and static equalization. Supported by our simulation, we found that the FM amplitude A_{pp} and FM frequency f_m of laser sources should be smaller than 1.5 MHz and 75 kHz, respectively, to avoid tracking failure. We identified an optimal parallelization level (P = 8), corresponding to a processing rate of 625 MHz. This rate is compatible with the state-of-the-art 40 nm CMOS technology [4], boding well for 64-QAM commercialization.

The above results are not only confined to 5 Gbaud systems. The feedback delay of parallel and pipelined DD-MLE is approximately P $\times$ d $\times$ T_{sym}, for P parallelization levels, d pipelining registers on the feedback path and T_{sym} symbol duration. Increasing the data rate from 5 to, say, 64 Gbaud, reduces T_{sym} to $T_{sym}/12.8$, while the parallelization level is required to increase by the same factor, i.e., from 8 to approximately 102. Therefore, the feedback delay of the phase tracking is unchanged and our conclusions are most likely applicable for higher than 5 Gbaud.

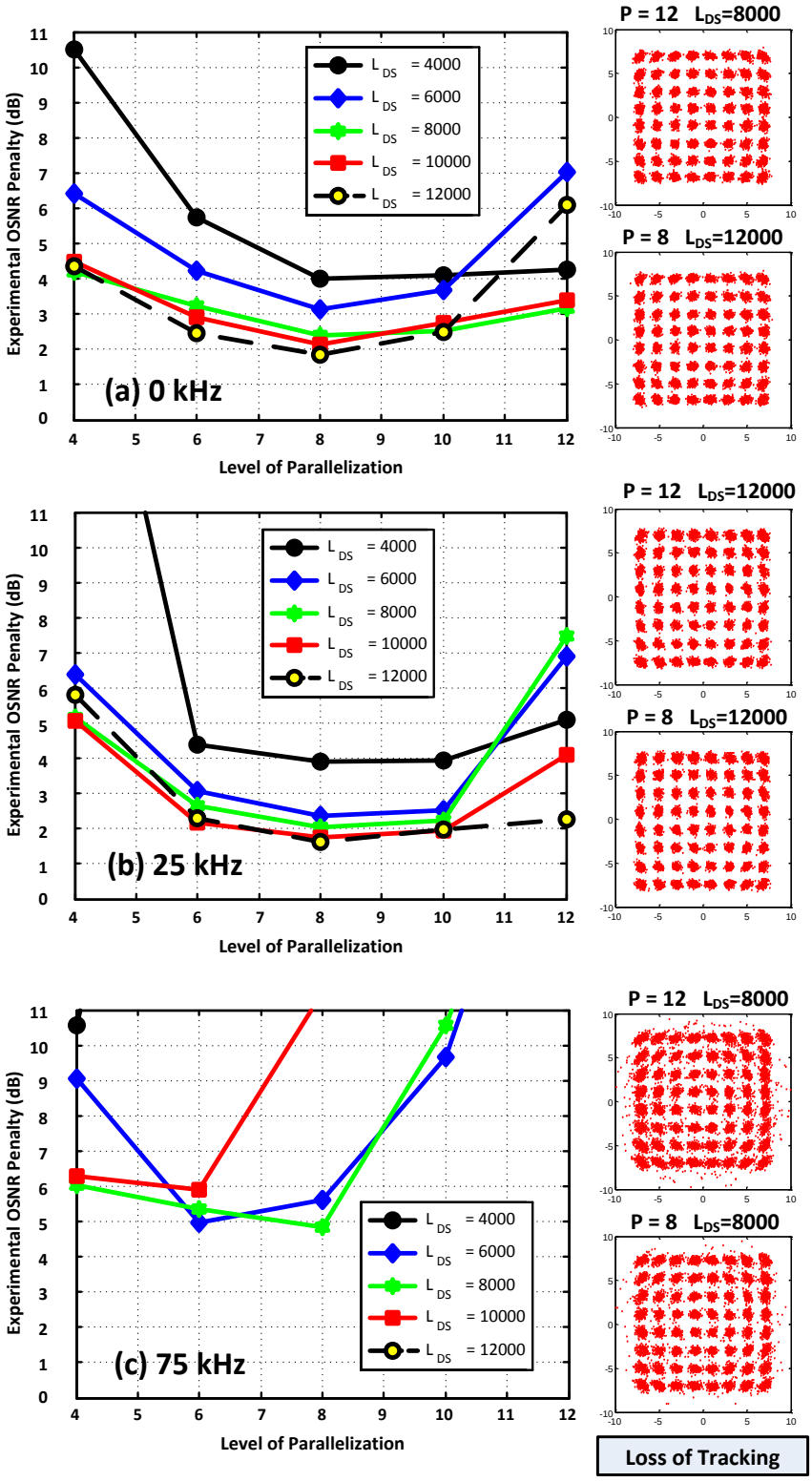

FIGURE 4.8 – Experimental OSNR penalty vs. levels of parallelization for (a) no FM, (b) sinusoidal FM, f_m = 25 kHz, A_{pp} = 0.6 MHz, (c) sinusoidal FM, f_m = 75 kHz, A_{pp} = 1.5 MHz. Right : the best constellations at received OSNR = 28 dB for each case.

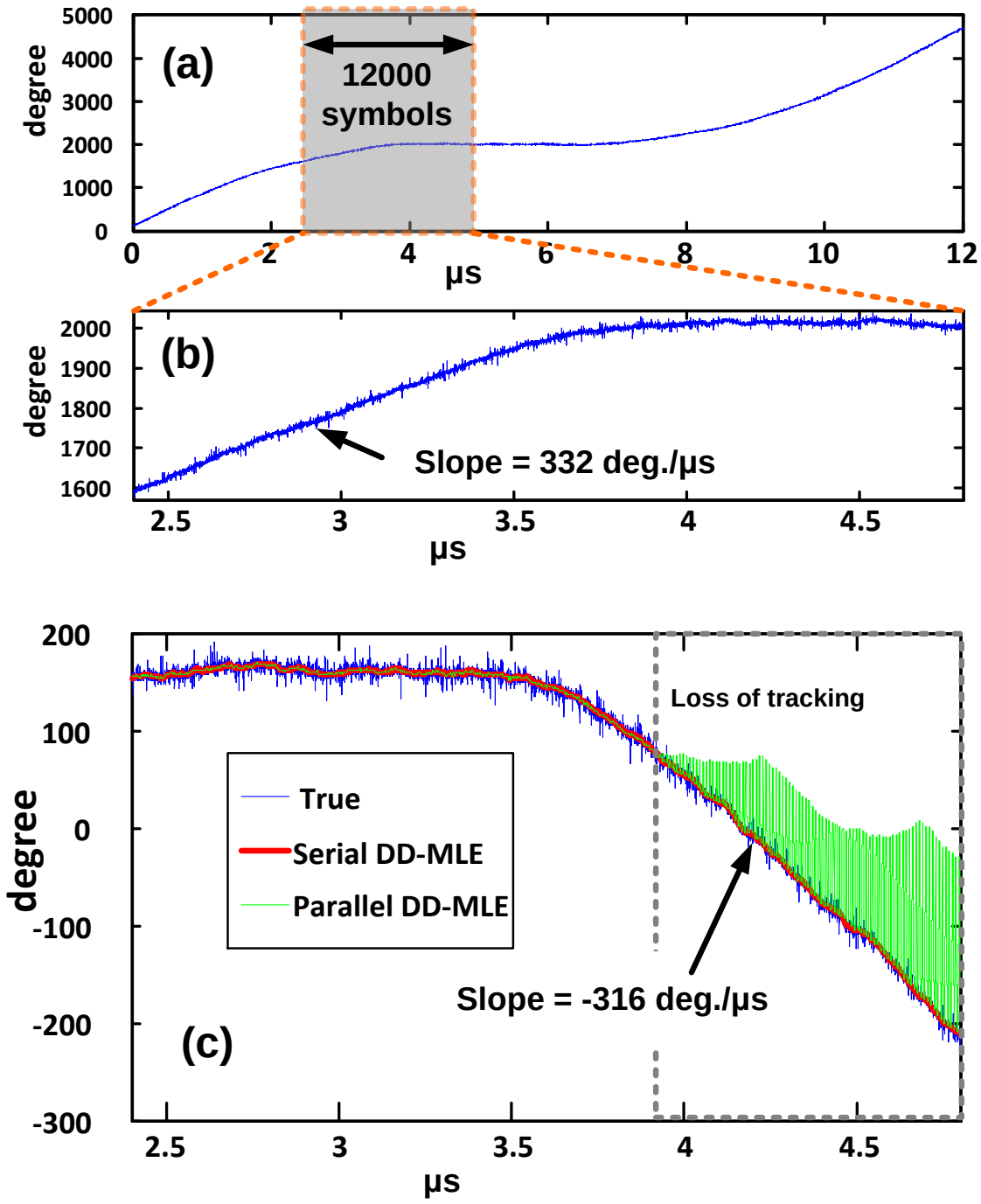

FIGURE 4.9 – (a) True laser phase with an FM of 75 kHz over 60 000 symbols (b) true laser phase noise from 12001th symbols to 24000th symbols (c) true phase, phase estimated by serial DD-MLE, phase estimated by parallel DD-MLE with P = 8.

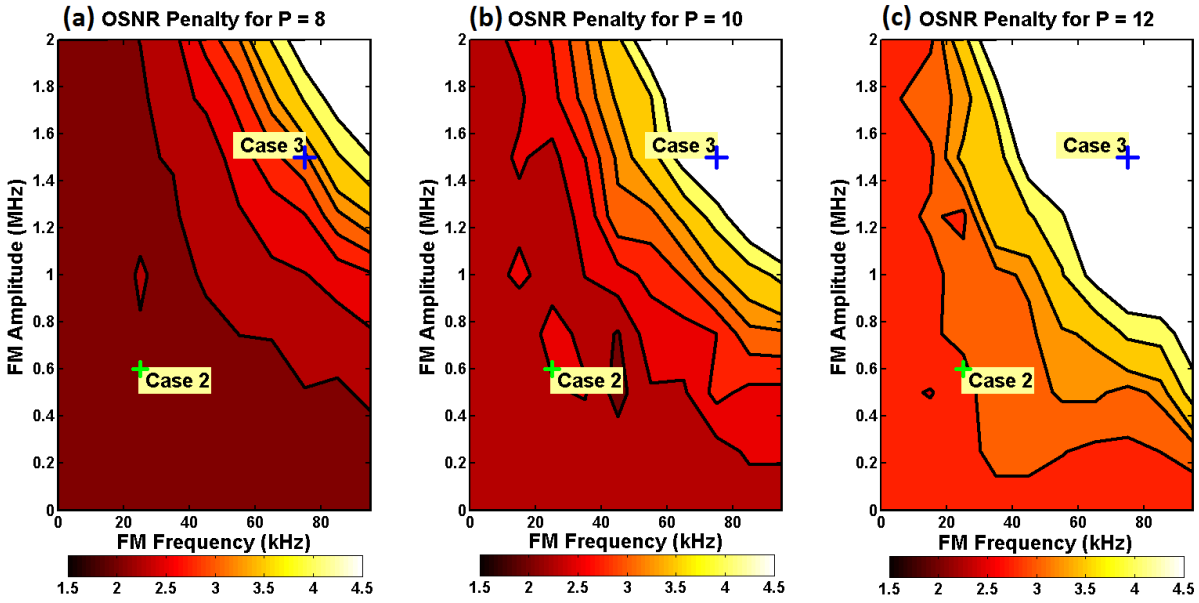

FIGURE 4.10 – OSNR penalty for sweeping f_m from 25 kHz to 75 kHz, A_{pp} from 0.5 MHz to 1.5 MHz. for (a) $P = 8$ (b) $P = 10$ (c) $P = 12$. All corresponds to LDS = 8k.

Chapitre 5

Digital Polarization Demultiplexing

In Chapters 2, 3 and 4, our main focus was on phase recovery in coherent receivers to understand the impact on system performance of laser frequency noise in different frequency regions. In the present chapter and after, we switch our attention from phase recovery to polarization demultiplexing. This chapter provides the background information for Chapter 6 and 7.

Spectral efficiency (information content within a frequency band) in fiber transmission can be doubled by using both X- and Y-polarizations to transmit two different data streams. Examples include the currently deployed DP-QPSK for 100G long-haul transmission. However, due to polarization rotation and polarization mode dispersion, the two polarization channels couple to one other, and the polarization crosstalk degrades the system performance. Polarization demultiplexing in coherent receivers is required to decouple the two polarization channels after fiber transmission.

In the present chapter, we will first give a complete overview of a DSP architecture of dual-polarization single-carrier digital coherent receiver. Next, we will discuss the model for fiber propagation polarization effects. Then, we will formulate the conventional MIMO (multi-input multi-output) system for polarization demultiplexing, and discuss the roles of MIMO. Problems of the conventional digital approaches will be explained. The solutions for these problems will be proposed in Chapter 6 and 7.

5.1 DSP Blocks in Dual-Polarization Single-Carrier Digital Coherent Receiver

In Chapters 2, 3 and 4, our main focus was on phase recovery in a single-polarization system. A single-polarization system helps us know clearly the impact of laser frequency noise on real-time phase tracking. From now on, we will focus on problems and solutions for digital polarization demultiplexing. Therefore, we will extend our discussion from a single-polarization digital coherent receiver in Fig. 2.1 to a dual-polarization digital coherent receiver in Fig. 5.1.

Some function blocks appearing in Fig. 2.1 are discussed again for completeness.

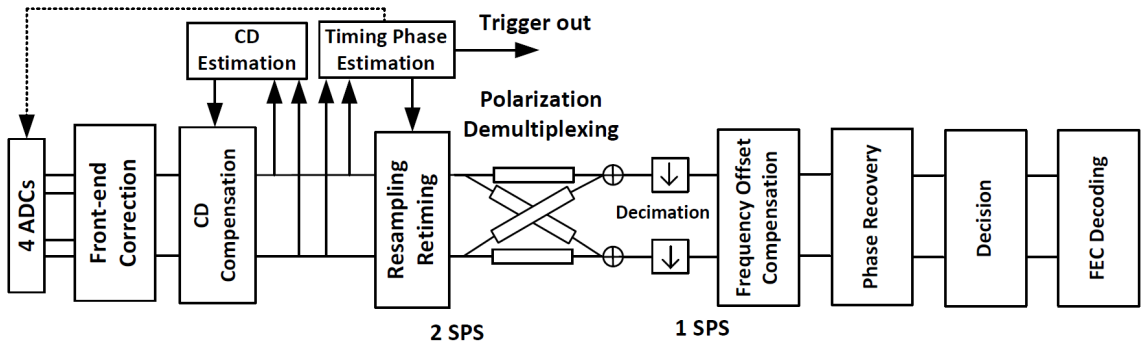

FIGURE 5.1 – The main blocks of digital signal processing in a single-carrier coherent receivers

1. **Analog-to-digital conversion (ADC)** : A dual-polarization optical signal (with advanced modulation like QPSK or QAM) is detected by a coherent receiver. The two polarization channels, with I and Q on each channel, correspond to four electrical analog signals at the output of the coherent receiver. The four ADCs perform sampling to digitize the four analog electrical signals, at a sampling rate defined by system and DSP designers, which is typically two samples per symbol (SPS). A lower sampling rate helps power saving, while a sampling rate of lower than 2×symbol rate introduces distortion (loss of information). The digitized signal contains the quantization error. Besides, due to the limited bandwidth of electrical components, the received signals are distorted by intersymbol interference.
2. **Front-end correction** : Timing skews between the four ADC channels must be corrected, in addition to IQ imbalance of the transmitter (critical for higher order modulation format).
3. **Chromatic dispersion compensation (CDC)** : It must be performed before the rest of DSP. After transmission, the phases of received signals are heavily distorted by chromatic dispersion, hence timing phase estimation cannot be performed, and the rest of the circuit cannot be triggered. CDC should be performed in frequency domain (FD) to reduce the hardware complexity [29]. Frequency offset compensation can be performed in this stage to remove a very large frequency offset [1] that affects the following equalization stage.
4. **Chromatic dispersion estimation** : CD estimation must be included on chips because coherent receivers should be made to adapt to different optical channels rather than be tailor-made for a particular channel (from the viewpoint of cost-effective ASIC chip production). The accuracy (error within 680 ps/nm, corresponding to 40 km) and the tolerance of the CD-estimation technique to polarization effects are necessary, to avoid

1. A large frequency offset shifts the center frequency of the received baseband signal from the zero frequency ; the DSP is designed only for baseband.

large OSNR penalty and unstable performance (large outage probability discussed in Chapter 6). One of the common techniques is based on the clock tone magnitude [57].

5. **Timing phase estimation (TPE)** : After CDC, the received signals become more coherent [2], allowing us to extract the timing phase. ADC Sampling instants will be skewed in the presence of clock frequency or timing phase mismatch between transmitters and receivers. Timing phase estimation is necessary during resampling and retiming to correct the sampling instant. The estimate can also be fed back to the ADCs for adjusting the sampling instant. However, the polarization effects and the bulk CD weaken the clock tone magnitude and distort the true timing phase. For the details of timing phase estimation, please refer to Appendix G.
6. **Resampling and retiming** : The clock frequency (of electronics) between transmitters and receivers may not be the same. It is likely that the sampling rate is less than 2×symbol rate in order to reduce the overall power consumption (For example, in Chapter 6, a 32 Gbaud signal is sampled at 40 Gsamples/s). Interpolation is performed to upsample to 2-SPS for the subsequent polarization demultiplexing. The retiming is performed by using the sampling phase estimated by TPE.
7. **Polarization demultiplexing** : In time-domain polarization demultiplexing, one uses the MIMO approach with T/2-spaced finite-impulse-response filters, which requires a 2-SPS input signal. It is mainly to compensate the polarization effects of an optical channel. The functions of MIMO will be discussed in Section 5.4 in detail.
8. **Decimation from 2-SPS to 1-SPS** is necessary after polarization demultiplexing. This helps reduce power consumption. The frequency offset compensation is performed to compensate the frequency mismatch between the transmit and receive lasers (which rotates the signal constellation).
9. **Phase recovery** : The laser phase noise rotates the constellation randomly. Phase noise estimation algorithms are required. Please refer to Chapter 2.
10. **Hard decision** : Only hard decision is considered in this thesis for symbol detection.
11. **FEC decoding** : Channel coding is beyond the scope of this thesis. The bit-error rates mentioned in this thesis are all without channel coding.

5.2 Model of fiber polarization effects

After the introduction of the main DSP blocks in a dual-polarization digital coherent receiver, this section will give a simple model that represents polarization effects in an optical channel. As chromatic dispersion (CD) is a relatively time-invariant effect compared to polarization

2. The CD-distorted coherent signals are not coherent as if they are random noises, since their phases are heavily distorted by the fiber dispersion. After CDC, the signals become "coherent" again, meaning a higher information content can be obtained from the data.

effects, we compensate the bulk CD before MIMO processing as mentioned in the previous section. The polarization effect of the fiber channel can therefore be modeled, in the frequency domain, as :

$$H(f) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{j(2\pi f\tau+\phi)/2} & 0 \\ 0 & e^{-j(2\pi f\tau+\phi)/2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{5.1}$$

where θ is the polarization orientation angle, ϕ is the phase delay between the optical fields on X and Y polarizations, τ is the differential group delay (DGD) due to polarization mode dispersion (PMD) in single-mode fiber (SMF).

For polarization demultiplexing, one has to compensate the effect of the optical channel in (5.1). As Jones matrix (5.1) acts on the field, one has to known both amplitude and phase of the received signal. This is the reason we require coherent detection system to double the spectral efficiency by using dual-polarization transmitted signals.

5.3 Conventional MIMO

In this section, the conventional MIMO for polarization demultiplexing will be revisited with input from [25, 26, 27, 36]

After CD compensation, resampling is performed to give 2 SPS. Retiming can be skipped before MIMO because the FIR filters of MIMO provide retiming via its interpolation property. Let $\underline{X}_{in}$ and $\underline{Y}_{in}$ be the input signals of the MIMO corresponding to X and Y polarization channels, with two SPS. $\underline{X}_{in}$ and $\underline{Y}_{in}$ suffer from the polarization crosstalks due to polarization rotation and DGD from each other, represented by the Jones matrix in (5.1). Our goal is to equalize this Jones matrix.

Let $\underline{X}_{in}^{k}$ be a 2-SPS input sequence that is convolved with two FIR filters $\underline{h}_{xx}^{k}$ and $\underline{h}_{yx}^{k}$, where k is the time index (increase by one for every sampling duration). All vectors here have the length of N, where N is the FIR filter length (number of taps). Similarly, $\underline{Y}_{in}^{k}$ is an 2-SPS input sequence that is convolved with two FIR filters $\underline{h}_{yy}^{k}$ and $\underline{h}_{xy}^{k}$. The outputs of MIMO, $X_{out}(k)$ and $Y_{out}(k)$, for X- and Y-polarizations respectively, are :

$$X_{out} = \underline{h}_{xx}^{k} \cdot \underline{X}_{in}^{k} + \underline{h}_{xy}^{k} \cdot \underline{Y}_{in} \tag{5.2}$$

$$Y_{out} = \underline{h}_{yx}^{k} \cdot \underline{X}_{in}^{k} + \underline{h}_{yy}^{k} \cdot \underline{Y}_{in} \tag{5.3}$$

where

$$\underline{X}_{in}^{k} = [X_{in,k}, X_{in,k-1}, \ldots, X_{in,k-N+1}]^T \tag{5.4}$$
$$\underline{Y}_{in}^{k} = [Y_{in,k}, Y_{in,k-1}, \ldots, Y_{in,k-N+1}]^T \tag{5.5}$$
$$\underline{h}_{xx}^{k} = [h_{xx,0}^{k}, h_{xx,1}^{k}, \ldots, h_{xx,0}^{k}]^T \tag{5.6}$$
$$\underline{h}_{xy}^{k} = [h_{xy,0}^{k}, h_{xy,1}^{k}, \ldots, h_{xy,0}^{k}]^T \tag{5.7}$$
$$\underline{h}_{yx}^{k} = [h_{yx,0}^{k}, h_{yx,1}^{k}, \ldots, h_{yx,0}^{k}]^T \tag{5.8}$$
$$\underline{h}_{yy}^{k} = [h_{yy,0}^{k}, h_{yy,1}^{k}, \ldots, h_{yy,0}^{k}]^T \tag{5.9}$$

The above FIR filters are adaptively updated by error signals, via a set of stochastic gradient algorithm equations with step-size parameter μ :

$$\underline{h}_{ij}^{k+1} = \underline{h}_{ij}^{k} + \mu \frac{\partial}{\partial \underline{h}_{ij}^{k}} \nabla J \tag{5.10}$$

where i and j can be x or y, ∇J is the averaged error signal depending on the criterion used for equalization. Typically, a blind adaptive approach is used based on constant modulus algorithm (CMA), such that the averaged error signal, ∇J_{CMA}, is

$$\nabla J_{CMA} = E\left[(|\underline{A}_{out}^{k}|^2 - R)^2\right] \tag{5.11}$$

where E denotes the expectation taken over the received samples, R is the average power of the QAM signal, $\underline{A}^k$ is either an output vector on either X or Y, depending on which polarization channel is chosen to be the output of FIR filters.

In practice, instantaneous error signals are used instead of an averaged error signal because the averaging requires long memory depth, i.e., $\nabla J_{CMA}^{k} = (|A_{out}^{k}|^2 - R)^2$. The set of equations of the gradient algorithm becomes :

$$\underline{h}_{xx}^{k+1} = \underline{h}_{xx}^{k} + \mu e_x(k') X_{out}(k) \underline{X}_{in}^{k} \tag{5.12}$$
$$\underline{h}_{xy}^{k+1} = \underline{h}_{xy}^{k} + \mu e_x(k') X_{out}(k) \underline{Y}_{in}^{k} \tag{5.13}$$
$$\underline{h}_{yx}^{k+1} = \underline{h}_{yx}^{k} + \mu e_y(k') Y_{out}(k) \underline{X}_{in}^{k} \tag{5.14}$$
$$\underline{h}_{yy}^{k+1} = \underline{h}_{yy}^{k} + \mu e_y(k') Y_{out}(k) \underline{Y}_{in}^{k} \tag{5.15}$$

where

$$e_x(k') = |X_{out}(k')|^2 - R \tag{5.16}$$
$$e_y(k') = |Y_{out}(k')|^2 - R \tag{5.17}$$

are the instantaneous error signals for X- and Y-polarization channels, respectively. Note that the set of equations (5.12) is updated only at every two samples, such that $k' = k - \mathrm{mod}(k, 2)$ [3],

3. $\mathrm{mod}(k, 2)$ is defined as the remainder after division of k by 2.

while the FIR filters h_{ij}^{k} is convolved with the input sequence samplewise. In this way, the FIR filters provide the retiming ability by shifting and interpolating samples.

As the instantaneous error has to be calculated to update the tap coefficients, feedback loops exist in real-time system. However, as the polarization effects ([5]) are slow compared to symbol rate, MIMO performance degradation is negligible compared to real-time phase tracking, as mentioned in Chapter 3.

Numerous works [77, 76, 9] demonstrate that following convergence of CMA, switching the error function in (5.11) from CMA to decision-directed least-mean square (DD-LMS) can further help improve the BER performance. However, DD-LMS requires symbol decisions after the phase recovery, resulting in a much larger feedback delay (between MIMO, frequency offset compensation, phase recovery) for implementation. Symbol decisions are performed during or after phase recovery, meaning that the equalization of MIMO is phase sensitive (being affected by the symbol-wise phase noise). In this thesis, we do not consider DD-LMS as our error function in MIMO-FIRs.

5.4 Role of conventional MIMO

In the state-of-the-art 100 Gbps system, conventional MIMO compensates the polarization effects of optical channel. As polarization effects are slow compared to symbol rate or sampling rate, MIMO functions not only as an equalizer for polarization effect, but also other equalization roles. The samplewise MIMO accomplishes the following functions :

1. **Polarization derotation** : The butterfly structure consists of four branches representing the four entries of an inverse Jone matrix. The cross components, $\underline{h}_{xy}^{k}$ and $\underline{h}_{yx}^{k}$ are responsible for compensating the crosstalk between two polarization channels.
2. **ISI equalization** : Equalization of the intersymbol interference (ISI) caused by limited receiver bandwidth is an intrinsic function of FIR filters.
3. **Residual CD Equalization** : Chromatic dispersion (CD) estimation is performed before the start of DSP shown in Fig. 5.1. However, the best CD estimation precision for long-haul transmission is only 40 km (corresponding to a residual CD of 680 ps/nm) [57]. The time-domain FIR filters must be long enough [50, 8] to compensate the residual CD to avoid large power penalty or a large system outage.
4. **Retiming** : Timing phase error due to the incorrect sampling instant introduced during analog-to-digital conversion [26] must be corrected before equalization. Timing phase error results in faulty equalization even with perfect knowledge of a channel. For a perfectly estimated inverse Jones matrix, analog-to-digital converters (ADCs) may not sample at the optimal points. The intrinsic time delay and interpolation properties of FIR filters with T/2-spacing with at least 5 taps are essential to avoid the OSNR power penalty due to timing phase errors [26].

5. **PMD compensation** : The DGD introduces crosstalk between two polarizations. This impairment can be adaptively compensated by the FIR filters of the cross components of the butterfly structure. For long-haul systems, the largest tap weights are usually a few taps off from the center, indicating a DGD of a few symbols introduced during the fiber transmission [9, 36].

5.5 Problems of conventional MIMO

Although MIMO provides several functions, it has the following problems :

1. **Singularity problem** : since the CMA is a blind algorithm, both X- and Y-polarization channels may converge to the same output [27, 72] ;
2. **Long Convergence** : Convergence time may be long for certain SOPs and for longer FIRs. The SOP-dependent convergence time, called the idle time, becomes critical for acquisition in burst-mode receivers for short-reach applications [35, 66, 33, 6] ;
3. **Power Consumption** : FIR filters are a sample-wise operation, which requires parallelization and a lot of multipliers to guarantee the equalization at sampling rate. It has high hardware complexity and is one of major power consumption, after CDC and FEC in coherent receiver ASIC chips [58, 29] ;
4. **Convergence Monitoring** : Convergence monitoring is required in coherent receivers and increases ASIC chip areas.

5.6 Summary

In this chapter, an overview of the DSP architecture in dual-polarization single-carrier digital coherent receivers was given. MIMO is used to perform digital polarization demultiplexing, equivalent to multiplying the received sampled signal with a Jones matrix representing the inverse of the optical channel. We have mathematically formulated the MIMO FIR, in which the CMA is commonly chosen for FIR filter coefficient updates. Finally, the roles of MIMO were explained : polarization derotation, ISI equalization, residual CD compensation, retiming and PMD compensation. The problems of MIMO, including the singularity problem, long convergence, which leads to significant power consumption and chip area, will be solved by our proposed DSP in Chapter 6 and 7.

Chapitre 6

SOP Pre-rotation before MIMO

In Chapter 5, we have reviewed the background for digital polarization demultiplexing, and understood the functions and problems of MIMO. In the present chapter, we will propose a DSP technique to solve the problems of MIMO and reduce power consumption in short-reach coherent communications. First, we will give a background for our motivation. We will introduce our proposed DSP architecture, its principle, and its implementation. We will go through the setup and results for our 100 Gb/s DPQPSK experiment with limited (16 GHz) receiver bandwidth. We will present a tradeoff between hardware reduction and performance degradation in the presence of residual chromatic dispersion for short-reach applications.

6.1 Introduction

Power consumption is a concern

Commercial coherent receivers at 100 Gbit/s have been deployed for long-haul transmission. Short-reach applications at these rates can take advantage of coherent technology as component prices decline. For short reach, power consumption can be reduced in subsystems for chromatic dispersion compensation (CDC) and forward error correction (FEC) from 50 W to 3 W [29, 58]. After CDC and FEC, polarization demultiplexers with 2×2 adaptive multi-input multi-output (MIMO) half-symbol-spaced finite-impulse response (FIR) filters remain a major source of power consumption, due to the large number of multipliers used for parallelizing the filters [5, 29].

For short-reach applications, PMD is much smaller than that in long-haul transmissions, and the length of FIR filters can be reduced. On the other hand, in the absence of channel impairment, the minimum FIR filters with $T/2$-spacing with at least 5 taps [26] are essential to mitigate residual timing phase errors as discussed in Chapter 5 [1]. A larger residual CD

1. Although the FIR filters correct the timing phase error due to their intrinsic interpolation ability [26], in practice, timing phase estimation algorithms are applied before the MIMO-FIRs to avoid time-consuming

requires longer FIR filters.

Retiming fails due to polarization effects

Other than chip power consumption, another practical concern is how we can obtain a clock signal from the received sampled signal in order to trigger the rest of the chip. After transmission, the signal's phase is heavily distorted by the fiber chromatic dispersion (CD), and therefore the clock information totally disappears. Thus, CD compensation must be performed before other equalizations. As shown in Fig. 5.1, CD estimation must be performed before CD compensation. For short reach systems (less than 100 km) or for long-haul coherent systems with optical or digital CD compensation, CD is not strong enough to totally suppress the clock information. Polarization effects, namely, polarization rotation and polarization mode dispersion, become the major impairments leading to failure our timing phase estimation (TPE) [57, 19].

For certain states of polarization (SOPs), the clock tone of signals completely disappears and TPE fails for half-symbol delay, or half-baud differential group delay (DGD)[19, 20, 31]. Loss of clock can be circumvented by pre-rotating the SOP to avoid 45-degree SOPs [62], or by applying the maximization of clock tones produced by various test SOPs [59, 60]. These feedback approaches, however, involve sophisticated DSP algorithms at symbol rates or higher and substantial numbers of block-wise multipliers [20, 62, 59, 60, 71]. Thus, these solutions are power hungry and increase implementation complexity. They exploit SOP pre-rotation (PR) before MIMO only to obtain a better clock phase, **not to assist polarization demultiplexing**. In the subsequent DSP, MIMO **repeat** the task of polarization rotation for polarization demultiplexing. For more information about TPE and clock tone, please refer to Appendix G.

Our contributions

In this chapter, we limit our attention to short-reach transmission (below 100 km), where DGD is present but not prominent [2]. We propose a novel parallelizable DSP architecture, to perform a very coarse SOP pre-rotation before MIMO using an inverse Jones matrix based on the minimization of only a single Stokes parameter (s_1). This SOP pre-rotation can coarsely reject the polarization crosstalk before the subsequent polarization demultiplexing DSP. It offers several advantages compared to the conventional MIMO-FIR approaches :

1. It requires only a single Stokes parameter (s_1) instead of all three Stokes parameters used in metrological algorithms in Appendix H, greatly reducing the computational

MIMO convergence. The residual timing phase errors due to estimation inaccuracy are then corrected by MIMO-FIRs.

2. According to Chapter 1 in [1], the PMD parameter D_p is 0.1-1 ps/km$^{\frac{1}{2}}$. The corresponding DGD is calculated as $D_p\sqrt{L}$, where L is the fiber length. For 40 km SMF, the the DGD value should be below 0.1 $\times\sqrt{40} = 6.3$ ps. However, D_p is 0.05 ps/km$^{\frac{1}{2}}$ in commerical fibers, i.e., the (mean) DGD value is 0.3 ps [37].

effort ;

2. It avoids the problem of long MIMO convergence and singularity for certain SOPs ;
3. Thanks to the reduced polarization coupling, we are then at liberty to reduce the number of MIMO cross-taps, leading to a significant reduction in number of MIMO-FIR CMs.
4. The SOP pre-rotation also brings the benefit of restoring clock tones for TPE even before MIMO. The overhead for coarse SOP estimation is easily counter balanced by savings in MIMO complexity.

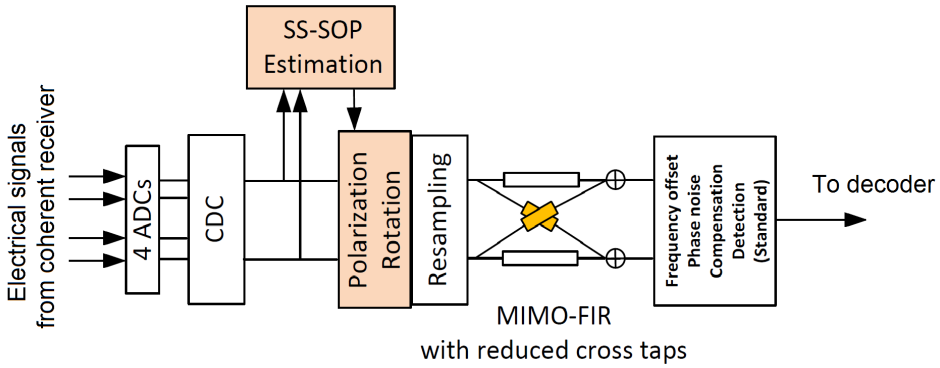

FIGURE 6.1 – DSP architecture with reduced cross-FIR taps. CDC : chromatic dispersion compensation. ADC : analog-to-digital conversion. SS-SOP : Stokes space state of polarization.

6.2 Proposal DSP Architecture

6.2.1 Overall Architecture

Fig. 6.1 shows our proposed DSP architecture. CDC is performed on the 4-channel sampled data. As SOP rotation (even in a long-haul system) is a slowly varying process compared to symbol rate, SOP estimation is only required at application-specific integrated circuit (ASIC) clock rates, which are below 500 MHz [50]. We propose our SOP estimation and SOP pre-rotation as shown in the two orange blocks in Fig. 6.1 : A very coarse SOP search in Stokes space (see Section 6.3) is performed continuously at every four ASIC clocks ; the number of ASIC clock cycles depends on the speed of SOP variation in the channel. The incoming data samples are pre-rotated at the sampling rate by an inverse Jones matrix using 12 real multiplies (RMs), a matrix calculated from the estimated SOP using Stokes parameter s_1. Finally, the interpolator resamples to two samples per symbol for the subsequent 2×2 MIMO $T/2$-spaced FIR filters. All other signal processing proceeds on the demultiplexed signals.

We propose to implement the MIMO-FIRs with reduced cross filter lengths (later we call it the reduced-complexity MIMO), as smaller FIRs are sufficient to compensate the small

TABLE 6.1 – Reduction in complex multipliers (CM) and the corresponding reduction percentage per ASIC clock period for various cross FIR lengths, N_{cross}, in our proposed reduced-complexity MIMO, compared to a full-complexity MIMO using $N_{cross} = 13$.

N_{cross}	CM reduction at 32 Gbaud	Reduction percentage
11	825	7.69 %
9	1704	15.38 %
7	2556	23.08 %
5	3408	30.77 %
3	4260	38.46 %

Remarks :
1 CM : complex multiplier.
2 ASIC rate : 300 MHz.
3 Assume that the straight FIRS have 13 taps.

residual polarization rotation. This results in a significant reduction in complex multipliers (CMs) : For optimal power consumption, ASIC clock rates run at **300 MHz** in current 28 nm CMOS technology [29]. The implementation of a parallel, $T/2$-spaced, N-tap FIR filter consists of $N \times P$ complex multipliers (CMs) to supply P outputs per ASIC clock [29], where P is the parallelization level. For example, a 13-tap $T/2$-spaced MIMO-FIR at 32 Gbaud requires $13{\times}P{\times}4 = 11076$ CMs per ASIC clock period, where the factor of 4 refers to the number of FIR filters in MIMO (two straight and two cross components). The parallelization level, P, is calculated as

$$P = \frac{R_s R_{os}}{R_{ASIC}}, \tag{6.1}$$

where R_s is the symbol rate, R_{os} is the oversampling ratio, defined as the number of samples per symbol in DSP, R_{ASIC} is the ASIC rate. For example, in our system, P is equal to $(32\text{G}{\times}2)/\mathbf{300M} \approx 213$ (the factor of two refers to the number of samples per symbol required for a $T/2$-spaced FIR). The total number of CMs used in our reduced-complexity MIMO is calculated as

$$N \times P \times 2 + N_{cross} \times P \times 2, \tag{6.2}$$

where N is the number of straight taps while N_{cross} is the number of cross taps. Our MIMO complexity reduction is calculated with respect to an original number of 13 taps, since the bit-error rate performance of the full-complexity MIMO did not show further improvement when using more than 13 taps [3]. Refer to Tab. 6.1, when the number of cross taps, N_{cross}, reduces from 13 to 11, the total number of CMs per ASIC clock period reduces from ($13{\times}P{\times}2$ + $13{\times}P{\times}2$) to ($13{\times}P{\times}2$+ $11{\times}P{\times}2$), i.e., a reduction of 852 CMs per ASIC clock period. The reduction percentage shown in Tab. 6.1 is defined as the number of CMs saved by shortening two cross FIRs divided by the total number of CMs used in a full-complexity MIMO, i.e.,

$$\frac{(N - N_{cross}) \times P \times 2}{N \times P \times 4} \times 100\%. \tag{6.3}$$

3. The optimization was performed on our measurement of 32 Gbaud DP-QPSK with a limited receiver bandwidth of 16 GHz.

Note that the previous calculation assumes the same complexity in the tap-update algorithm (CMA) in both full-complexity and reduced-complexity MIMOs.

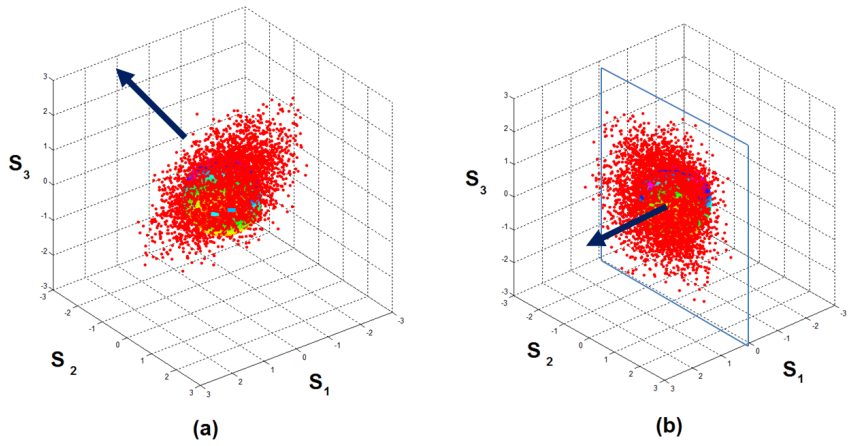

FIGURE 6.2 – Principle of SOP estimation based on s_1-MMSD. (a) Before polarization rotation ; (b) after polarization rotation according to S1-MMSD.

6.3 SOP-Search Implementation in Stokes Space

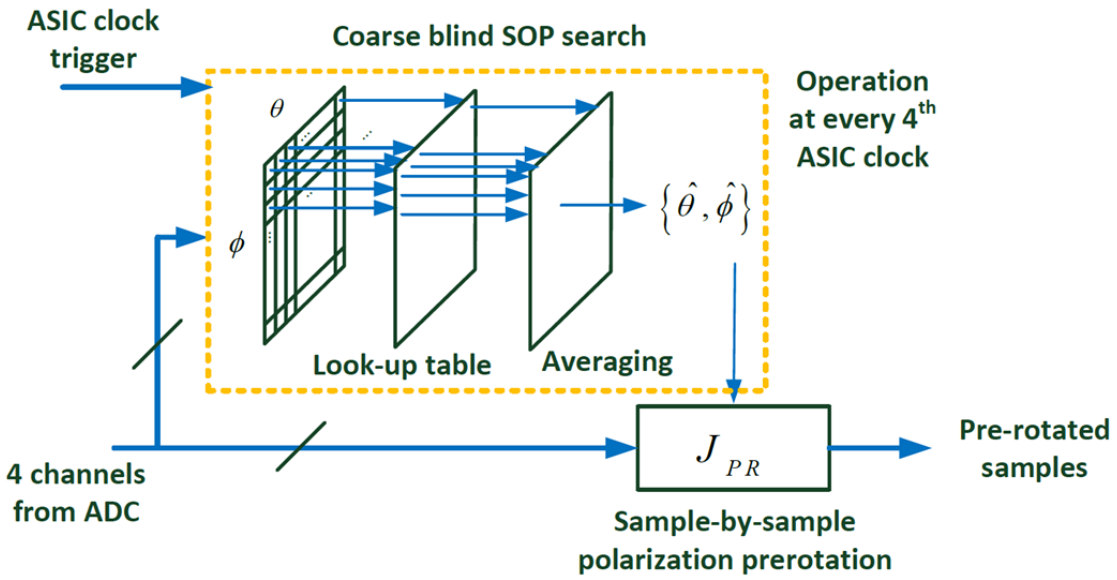

FIGURE 6.3 – Implementation of feedforward blind SOP search based on s_1 parameter at every 4th ASIC clock.

Szafraniec et al. [63, 64] and Muga et al. [38] applied a plane fitting and a classical adaptive tracking algorithm, respectively, to obtain an inverse Jones matrix based on the Stokes parameters of a normal to the lens-shaped object (Please refer to Appendix H or Section 3.2 in [64]) . After SOP rotation, the resultant signals have similar power on X and Y polarizations. These methods evaluated all three Stokes parameters, leading to significant increased complexity.

Instead of evaluating three Stokes parameters with plane fitting or tracking, we find the SOP estimate, $\hat{s}$, (i.e., the polarization angle θ and the phase delay ϕ between two polarizations) via the minimum mean squared distance (MMSD) between the received samples and the plane at $s_1 = 0$ as shown in Fig. 6.2, i.e.,

$$\hat{s} = \min_{s} E[|s_1|^2], \tag{6.4}$$

where s_1 is the first Stokes parameter [63] defined as

$$s_1 \triangleq |E_x|^2 - |E_y|^2, \tag{6.5}$$

with E_x and E_y representing the complex fields at the X- and Y-polarizations with respect to coherent receiver. As Stokes parameters are functions of Jones-space parameters, θ and ϕ, we will minimize $|s_1|^2$ in terms of θ and ϕ, instead of Stokes vector s. For convenience, from now on we will call "minimizing s_1" insteand of $|s_1|^2$. This results in a 2-D SOP blind search based on minimizing s_1, i.e., (6.4) is equivalent to the following :

$$\left\{\hat{\theta}, \hat{\phi}\right\} = \min_{\{\theta,\phi\}} E[|s_1|^2]. \tag{6.6}$$

First, the received complex signals in both polarizations are rotated by various Jones matrices in parallel (12 RMs) [51], as shown in the first stage of the dotted box in Fig. 6.3 :

$$J_{PR} = \begin{bmatrix} \cos\theta & -\sin\theta e^{i\phi} \\ \sin\theta & \cos\theta e^{i\phi} \end{bmatrix} \tag{6.7}$$

where the negative sign acts as a constraint to avoid the non-physical singularity that troubles the constant modulus algorithm [27] [4]. We use only 10° resolution for both θ (spanning 90°) and ϕ (spanning 180°), which requires 18×9×12 = 1944 RMs or 486 CMs each time.

4. The unitary representation of the Jones matrix acts as a natural constraint to prevent the non-physical singularity, where singularity refers to the outcome that two polarization channels at the output converge to the same polarization channel data [27]. To see that our J_{PR} is unitary, (7.14) can be written as

$$J_{PR} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{bmatrix}. \tag{6.8}$$

The term $e^{i\phi}$ represents the phase delay between x and y polarizations. Since Jones representation depends on the relative polarization angle and phase between x and y polarizations, a common phase factor $e^{-i\phi/2}$ can be introduced to x and y polarizations in (6.9) :

$$J_{PR} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{bmatrix} = \begin{bmatrix} \cos\theta e^{-i\phi/2} & -\sin\theta e^{-i\phi/2} \\ \sin\theta e^{i\phi/2} & \cos\theta e^{i\phi/2} \end{bmatrix}. \tag{6.9}$$

The determinant of J_{PR} is one.

As the SOP rotation is slow, we need only acquire samples and perform the SOP estimation at every four ASIC clocks, i.e., 121.5 CMs per clock. This increase in CMs is compensated by the reduced number of CMs in our modified MIMO-FIRs shown in Tab. 6.1. The $18\times9 = 162$ rotated complex fields will be converted to the corresponding s_1 (and thus $|s_1|^2$) using look-up tables. The result is saved into registers and added to the next s_1 values. An average is taken over a block of 200 samples (time-interleaved at every four ASIC clock periods) to generate the mean squared distance (MSD). The optimal SOP corresponding to MMSD is chosen by comparator circuitry. The two-dimensional MSD surface is shown in Fig. 6.4.

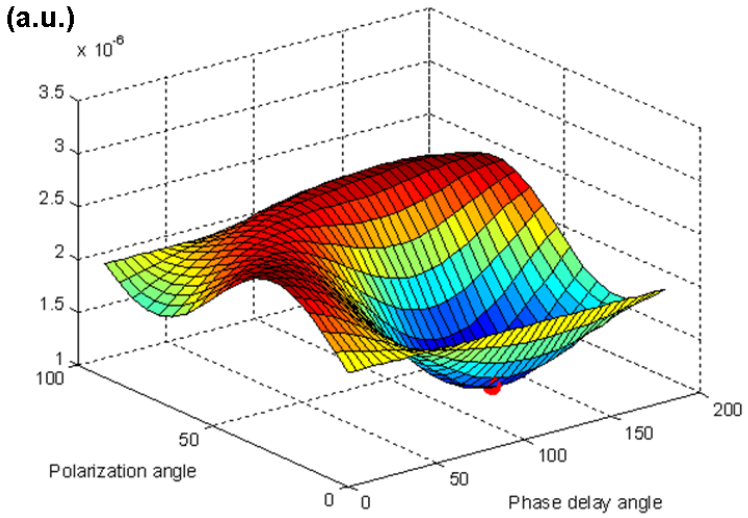

FIGURE 6.4 – The surface of the mean squared distance (MSD) of S_1 parameter. Red dot refers to the MMSD (the optimal SOP).

Finally, we rotate every sample using 12 RMs using the optimal SOP, requiring $3\times$ (f_s/300 MHz) CMs per ASIC clock period, where f_s is the sampling frequency, as shown in the lower part of Fig. 6.3. In our work, f_s is 40 GHz, and the total number of the required CMs per clock period is $400 + 121.5 = 521.5$. Tab. 6.1 shows that, for 32 Gbaud DP-QPSK, the overall CM reduction starts when reducing N_{cross} by more than two. Obviously, the computation complexity can be further reduced by using cascaded blind search.

6.4 Experimental Setup

The performance of our proposed DSP is examined using the experimental setup shown in Fig. 6.5. A 32 Gbaud DPQPSK signal was generated by a polarization multiplexing emulator (dashed, orange), with laser sources of 10 kHz linewidth. An Agilent N7786B polarization synthesizer was used to change the SOP of the transmitted signal. After passing through an erbium doped fiber amplifier (EDFA) and 40 km standard single-mode fiber (SMF) fiber, the

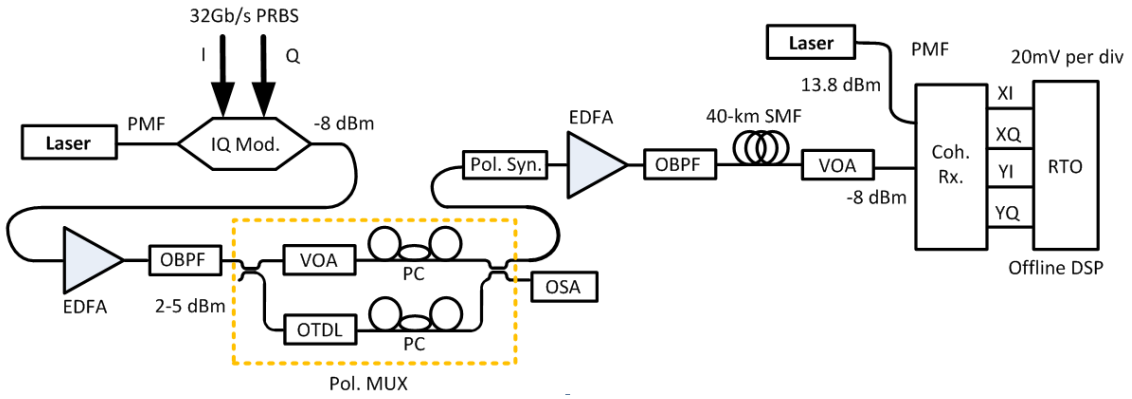

FIGURE 6.5 – Experimental setup for 32 Gbaud DP-QPSK 40 km transmission. PC : Polarization controller, IQ Mod. : Inphase-Quadrature modulator, Pol. Syn. : Polarization synthesis, OBPF : optical band pass filter, OTDL : Optical tunable delay line, OSA : optical spectrum analyser, RTO : real-time oscilloscope, EDFA : erbium doped fiber amplifier, VOA : variable optical attenuator, Coh. Rx. : Coherent receiver.

received signal was coherently detected by an integrated coherent receiver having a bandwidth of 22 GHz. The real-time oscilloscope (ADC) sampled at 40 GSa/s with bandwidth of 16 GHz. Offline processing was used and 1,120,000 bits were used for bit error rate (BER) measurement.

TABLE 6.2 – Cases considered for BER performance analysis.

Case	SOP-PR	N_{cross}	CMs for PR	CMs for MIMO	Acq. time	CDC before MIMO
X	No	13	None	11076	long	low
A	Yes	13	None	11076	short	low
B	Yes	7, 5	521.5	8520, 7668	short	medium
C	No	7, 5	521.5	8520, 7668	long	medium

Remarks :
1 SOP : state-of-polarization ;
2 PR : pre-rotation ;
3 CM : complex multipliers (per ASIC clock period) ;
4 Acq. : acquisition ;
5 CDC : chromatic dispersion compensation.

For DSP, the flow shown in Fig. 6.1 was used. CDC was first performed on the captured samples. Then a conventional MIMO approach and our proposed DSP for polarization demultiplexing were performed. Tab. 6.2 shows four cases of interest in this chapter, assuming the same parameters (i.e. 32 Gbaud, 13 straight taps, variable cross tap numbers) as Tab. 6.1. In all cases, the CMA algorithm was employed for calculating the errors and updating the FIR taps as explained in Chapter 5. After polarization demultiplexing, standard algorithms for frequency offset compensation and carrier phase recovery were performed, followed by detection and bit error counting.

Case X refers to the conventional MIMO-FIR, where the acquisition time can be long for certain SOPs. Case A refers to the conventional MIMO-FIR with our proposed SOP estimation/pre-rotation (PR), which reduces the acquisition time but otherwise leaves performance unchanged [5]. Case B refers to **our approach**, in which we have our proposed SOP-estimation/SOP-PR and we reduce the MIMO-FIR cross taps from 13 to 7 and 5, respectively. Case C is similar to Case B, but without SOP estimation/SOP-PR. Case C is a control to allow us to understand the importance of SOP-PR for reduced-complexity MIMO. As Case X and Case A make use of more cross taps, they are expected to have a better CD tolerance and requires less effort for CDC before MIMO compared to Case B and Case C. We do not repeat case X, since Case A will share the same BER performance result, except that it requires a longer acquisition time.
Tap initialization is important in both real-time and offline DSP. In the state-of-the-art digital coherent receivers, training sequences are employed to reduce the convergence time of tap values [45, 46], sacrificing transmission capacity. In offline DSP, convergence time of the CMA algorithm should be minimized, so that we can calculate BER faster. One may expect that the reduced-complexity MIMO should converge faster than the full-complexity MIMO because of shorter tap length. However, based on our finding, the "overly" reduced-complexity MIMO (i.e., $N_{cross} = 5$) cannot converge at all because of the large error signal (the tap updates via CMA depends on the error signals).

In our offline DSP, the full-complexity MIMO (with $N_{cross} = 13$) in case A requires no tap initialization. The tap values can finally converge, and the error signal stays small after convergence because the FIR taps are long enough for equalization. However, singularity may occur. In our measurement, SOP estimation/PR was performed to speed up the convergence (and avoid singularity) for saving our processing time. In contrast, the reduced-complexity MIMO in case B and case C should require tap initialization, since inadequate numbers of cross taps result in partial equalization, and thus larger error signals and fluctuation, leading to erroneous tap-value calculations and failure of convergence. This tap initialization was replaced by our proposed SOP estimation/PR, which compensates coarsely the polarization effects, reducing the effort in tap convergence, as shown in the blue curves in Fig. 6.6. However, without SOP estimation/PR, tap initialization was performed by using the information from the best case (case A ; full-complexity MIMO) after all taps converge, as shown in the black curves in Fig. 6.6. We employed the SOP pre-knowledge only for achieving the best performance for case C (without SOP estimation/PR), i.e., the BER performance of case C shown in Fig. 6.6 is already the best that the reduced-complexity MIMO can achieve.

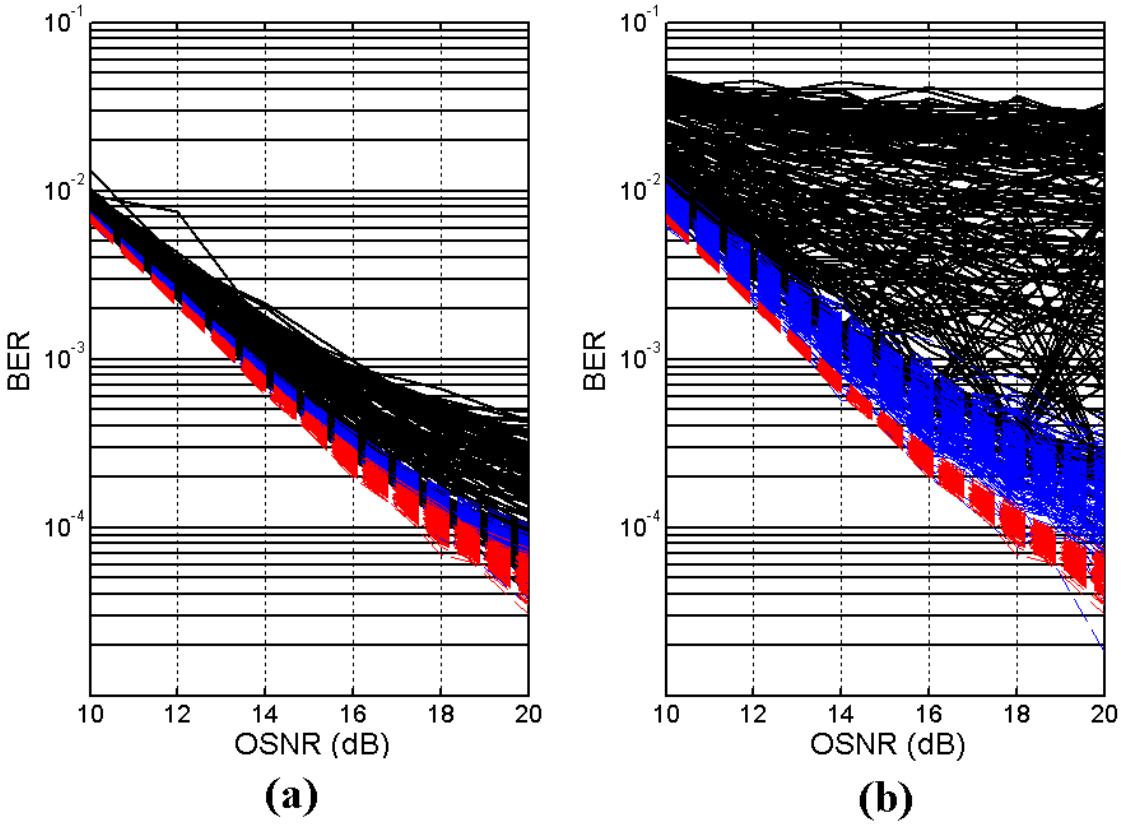

FIGURE 6.6 – Comparison in BER performance between the conventional MIMO-FIR (red), our approach with SOP estimation/PR (blue) and the reduced-complexity MIMO without SOP estimation/PR (black) in the presence of 40 km-SMF transmission (a) Black : 7 cross taps without SOP prerotation ; with tap initialization*. Blue : 7 cross taps with SOP prerotation. Red : 13 cross taps with SOP prerotation. (b) Black : 5 cross taps without SOP prerotation ; with tap initialization*. Blue : 5 cross taps with SOP prerotation. Red : 13 cross taps with SOP prerotation. All above were generated using 10 degree-resolution blind SOP search.(*) Without SOP prerotation, MIMO convergence must be assisted with prior information of SOP. This prior SOP can be obtained from the best case (case A ; full-complexity MIMO) after all taps converge, and is then used to initialize the FIR taps for case C (reduced-complexity MIMO without SOP-PR).

6.5 BER performance for various SOPs

To investigate the robustness of our proposed DSP, we performed an experiment to test the DSP with random SOPs covering the entire Poincaré sphere (please refer to Appendix I for experimental detail). For each SOP, we swept the OSNR from 10 dB to 20 dB with a 2 dB step. For each OSNR value, we captured 10 realizsations (data) of 10k samples to

5. Note that our proposed SOP estimation/PR adds extra computational complexity to the full-complexity MIMO ; Case B compensates this additional complexity to take cross FIR taps away from the MIMO structure.

calculate the BER values. Fig. 6.6 shows the comparison in BER versus optical signal-to-noise ratio (OSNR) between a full-complexity MIMO-FIR (Case A ; red), our proposed reduced-complexity MIMO-FIR with SOP-PR (Case B ; blue), and the reduced MIMO-FIR without SOP-PR (Case C ; black). Each BER curve represents a unique SOP. In total, there are 245 unique SOPs uniformly distributed on the Poincaré sphere. Note that individual BER curves were plotted instead of taking average over all curves, as our intention is to show the robustness of various DSP methods, i.e., fluctuation of BER curves over various SOPs. We consider two levels of complexity reduction. Fig. 6.6a and Fig. 6.6b correspond to N_{cross} = 7, and 5, respectively. The full-complexity MIMO-FIR (Case A ; red) with a SOP-PR stage included to assure good convergence for all SOPs, is shown in both figures for comparison.

Fig. 6.6b refers to case C (without SOP-PR ; black ; $N_{cross} = 5$). The BER curves for various SOPs consist of discontinuities, showing that the algorithm never converged for some SNR points, and therefore the BER plummeted.

As shown in Fig. 6.6a and Fig. 6.6b, our new approach with SOP-PR (case B ; blue) can reduce the number of cross taps and maintain performance similar to the conventional approach (case A ; red). Our experimental results show that case C (without SOP-PR ; black) yields a higher power penalty than that of our proposed DSP in case B (with SOP-PR). The variation in BER at a fixed OSNR is significant when reducing MIMO complexity without providing an initial de-rotation to a benign initial SOP. The performance-complexity (N_{cross}) trade-off is visible when comparing Fig. 6.6a (greater complexity, improved performance) and Fig. 6.6b (reduced complexity, but worse performance).

In Fig. 6.7 we examine the statistics for MIMO-FIR values. The means and variances of the absolute value of the filter taps are shown for all three cases. Results are shown for the both scenarios of Fig. 6.6. Statistics were calculated by averaging over 245 unique SOPs, with 10 realizations for each of the 245 SOPs. In total, the value at each point in Fig. 6.6 was calculated by taking mean or variance over 2450 realizations after tap convergence.

Fig. 6.7a, the full-complexity approach (case A ; red squared), shows that the assistance of SOP-PR can effectively shrink the mean cross-tap values below 0.1. This shows one of the advantages of SOP-PR : N_{cross} could be reduced to simplify the hardware with minimal performance degradation by means of SOP-PR. Our proposed DSP with SOP-PR (case B, blue cross) reduces N_{cross} from 13 to 7, gives mean cross-tap values below 0.25, resulting in negligible OSNR penalty. However, without SOP-PR, the mean cross-tap values can be higher than 0.5 (case C ; black circle), resulting in higher penalty shown in Fig. 6.6a. This implies that cross taps with significant values must be used to cancel the effect of polarization crosstalk.

The scenario of $N_{cross} = 5$, corresponding to the BER performance in Fig. 6.6b, shows our proposed DSP (case B, with SOP-PR ; blue) can maintain similar performance to the case of

$N_{cross} = 7$, allowing further reduction in CMs. Case C (without SOP-PR ; black) does not work. Refer to Fig. 6.6b, one may argue that the vigorous BER fluctuation could be due to the problems of our DSP algorithms. A better BER calculation should be obtained by paying more effort on monitoring MIMO convergence in our DSP.

The BER fluctuation (discontinuities) and high OSNR penalty of case C could be a consequence of inadequate number of cross taps ; it is not caused by the problem in monitoring convergence. The same CMA algorithm was applied to all cases, and we had already initialize the FIR taps in Case C by using the polarization information from case A (full-complexity MIMO). Fig. 6.6b shows that case C, without SOP-PR but with large tap reductions, does not work at all with conventional standard DSP. We aim at showing the importance or necessity of using SOP estimation/PR for the reduced-complexity MIMO.

In Fig. 6.7b, the mean cross-tap values of our proposed DSP (case B ; blue squared) are all below 0.3, hence the negligible OSNR penalty. However, case C (without SOP-PR) has mean cross-tap values higher than 0.9. This is because CMA requires only the error signal from the FIR filters for calculating the tap values. Inadequate numbers of cross taps result in partial equalization and thus larger error signals and fluctuation, leading to erroneous tap-value calculations even in the presence of SOP pre-knowledge.

We next examine the required precision for our blind SOP search. Could further precision in SOP estimation lead to even greater savings in MIMO-FIR complexity ? Fig. 6.8 shows the comparison between 10 degree and 1 degree SOP search resolution for the case of $N_{cross} = 5$ after 40-km full CDC. Surprisingly, 10 degree resolution is already sufficient for maintaining BER performance similar to that of 1-degree resolution. We have not tried a detailed optimization in terms of resolution and penalty, as we only address the novelty of our proposed DSP. Note that a very high SOP estimation is not necessary, as a pure SOP-PR using an inverse Jones matrix as in Stokes-space PolDemux (refer to our discussion in Appendix H), no matter how precise the resolution can be, is insufficient for achieving our performance. As discussed previously, the presence of cross-taps is indispensable for compensating residual polarization-crosstalk and small DGD compensation and for retiming through the FIR interpolation ability.

6.5.1 CD tolerance

CDC is an important source of power consumption in an ASIC chip [58, 29]. The ability of the MIMO-FIR to compensate residual CD should not be compromised, so that we can maintain CDC power consumption at the lowest possible level. In short-reach scenarios, CD exists and must be compensated, but a separate DSP stage for CD compensation is not preferred. The reason is pragmatic and cost-effective : OSNR is much higher than the OSNR threshold for FEC within 100 km transmission, and therefore system providers need only control the power

penalty due to residual CD within the link budget. In this section, we extend our analysis to examine tolerance to residual CD of our proposed DSP. Fig. 6.9a and Fig. 6.9b show the probability mass functions (pmfs) of BER conditioned on different OSNRs values of cases A, B and C, which allow us to analyze the outage probability (Please refer to Section 6.7) in the presence of residual CD for the case of $N_{cross} = 7$ and the case of $N_{cross} = 5$, respectively. We used the experimental set-up of 40 km transmission in Fig. 6.5. The CDC (shown in Fig. 6.9) was set to compensate only 0, 20 and 40 km (680, 340 and 0 ps/nm residual CD, respectively) shown in green, red and blue, respectively. Although CD should be constant in our 40 km SMF, it becomes stochastic when the polarization crosstalk (which depends on the SOP distribution) can be not compensated perfectly by a reduced-complexity MIMO.

Generally, the mean of the pmf of BER shifts to higher values and the variance of the pmf of BER becomes larger when the residual CD increases. For example, considering $N_{cross} = 7$ in Fig. 6.9a, our proposed DSP allows a 20 km short reach transmission (without CDC) at OSNR = 18 dB [6] with low system outage (a large portion of the pmf of BER is below log(BER) = -3, the FEC level). This tells us how N_{cross} should be to trim our coherent DSP complexity for a certain performance in short reach applications.

6.5.2 Restoration of Clock Tone Magnitude

In practice, after CDC, we resample the signal to 2 samples per symbol, which requires us to determine the sampling phase using frequency-domain (FD) Godard timing phase estimation (TPE) [20, 62, 59, 60, 60, 71]. The principle of FD-Godard TPE is to obtain the phase of the autocorrelation function of the received signal (2 samples per symbol) spectrum at the clock frequency. The complex value of this autocorrelation function at the clock frequency is defined as our complex-valued clock tone [57]. For more details, please refer to Appendix G.

The clock tone magnitude should be large enough to provide a good S-curve phase detector for timing phase recovery circuit [20, 62, 59, 60, 71]. It is shown in [22] that FD-Godard TPE is equivalent to the time-domain square-law nonlinearity TPE [41] which requires 4 samples per symbol. One prefers to perform TPE and resampling in the frequency domain to minimize the hardware complexity [57].

The previous TPE algorithms [59, 60, 61, 62] require SOP-PR to avoid clock tone loss due to polarization effects. The advantage of using our proposed DSP is shown in Fig. 6.10, in which the clock tone magnitudes were obtained using FD-Godard TPE [57] over 780 various unrepeated SOPs distributed over the Poincaré sphere. The conventional approach without SOP-PR (red) suffers from clock tone loss for certain SOPs (under small DGDs, polarization dependent loss, timing offset between X and Y optical signals, and different sampling offsets between X and Y analog-to-digital converters), leading to system failure of coherent receivers.

6. 18 dB or above are attainable OSNR values for shor t-reach systems.

Our proposed DSP with SOP-PR (black) can restore the weakened clock tone magnitude. Note that in our experiment, the DGD introduced by the 40-km fiber span should be negligible. We introduced a fixed timing offset between X and Y at the transmitter using an optical tunable delay line (OTDL) in Fig. 6.5 to emulate the loss of CT shown in Fig. 6.10 (red).

6.6 Conclusions

In this chapter, we demonstrated that an SOP pre-rotation (by multiplying the received samples with an estimated Jones matrix corresponding to the inverse of an optical channel) can be performed before a conventional 2×2 MIMO for polarization demultiplexing. Its advantages are to avoid long convergence of the conventional MIMO and singularity problems at certain SOPs, suitable for applications requiring fast acquisition such as burst-mode receivers in packet-switching-based short-reach or metro network. The inverse Jones matrix requires us to estimate the SOP, i.e., we generate our own channel-state information (CSI) before polarization demultiplexing is performed.

Firstly, instead of estimating SOPs using all three Stokes parameters (s_1, s_2, s_3) with complicated plane fitting previously proposed for metrological purpose, we have proposed a very coarse SOP estimation based on minimizing a single Stokes parameter, s_1, to reduce the computational effort, with a 10-degree resolution only. Secondly, as this SOP pre-rotation can coarsely reject the polarization crosstalk before the subsequent polarization demultiplexing, i.e. the polarization coupling is reduced, we are then at liberty to reduce the number of MIMO cross-taps, leading to a significant reduction in number of MIMO-FIR complex multiplications, i.e., the overhead for coarse SOP estimation is easily counter balanced by savings in MIMO complexity. Finally, the SOP pre-rotation also brings the benefit of restoring clock tones for TPE even before PolDemux.

Our proposed DSP was validated by our experimental results in a 100 Gb/s DPQPSK system with limited (16 GHz) receiver bandwidth. We have also investigated a tradeoff between hardware reduction and performance degradation in the presence of residual chromatic dispersion for short-reach applications.

6.7 Outpage Probability

The definition of outage probability in this work is given here. The values of BER of a 2×2 MIMO-FIRs (CMA-based) with N straight taps and N_{cross} taps follows a probability density function (pdf) conditioned on a given SOP (state-of-polarization) s, i.e. $p_{N,N_{cross}}(\gamma|s)$, where the SOP is described by azimuth angle θ (polarization rotation) and ellipticity angle ϕ (phase delay between two polarizations). The outage probability of a 2×2 MIMO-FIRs (CMA-based) with N straight taps and N_{cross} taps is defined as the probability of which at a given OSNR,

the averaged BER is above a certain BER value γ, for example, FEC, i.e.

$$OP_{N,N_{cross}}(FEC, OSNR) = \int_{\gamma > FEC} \int_{s} p_{N,N_{cross}}(\gamma|s) f_s(s) ds d\gamma \tag{6.10}$$

where the pdf of SOPs $f_s(s)$ is assumed to be uniformly distributed over the Poincare sphere.Due to the limited number of realizations taken from our measurement, the corresponding probably mass functions (pmfs), $P_{N,N_{cross}}[\gamma_j|s_i]$ and $P_s[s_i]$ are used instead of a continuous pdf :

$$OP_{N,N_{cross}}(FEC, OSNR) \approx \sum_{\gamma_j > FEC} \sum_{s_i} P_{N,N_{cross}}[\gamma_j|s_i] P_s[s_i] \tag{6.11}$$

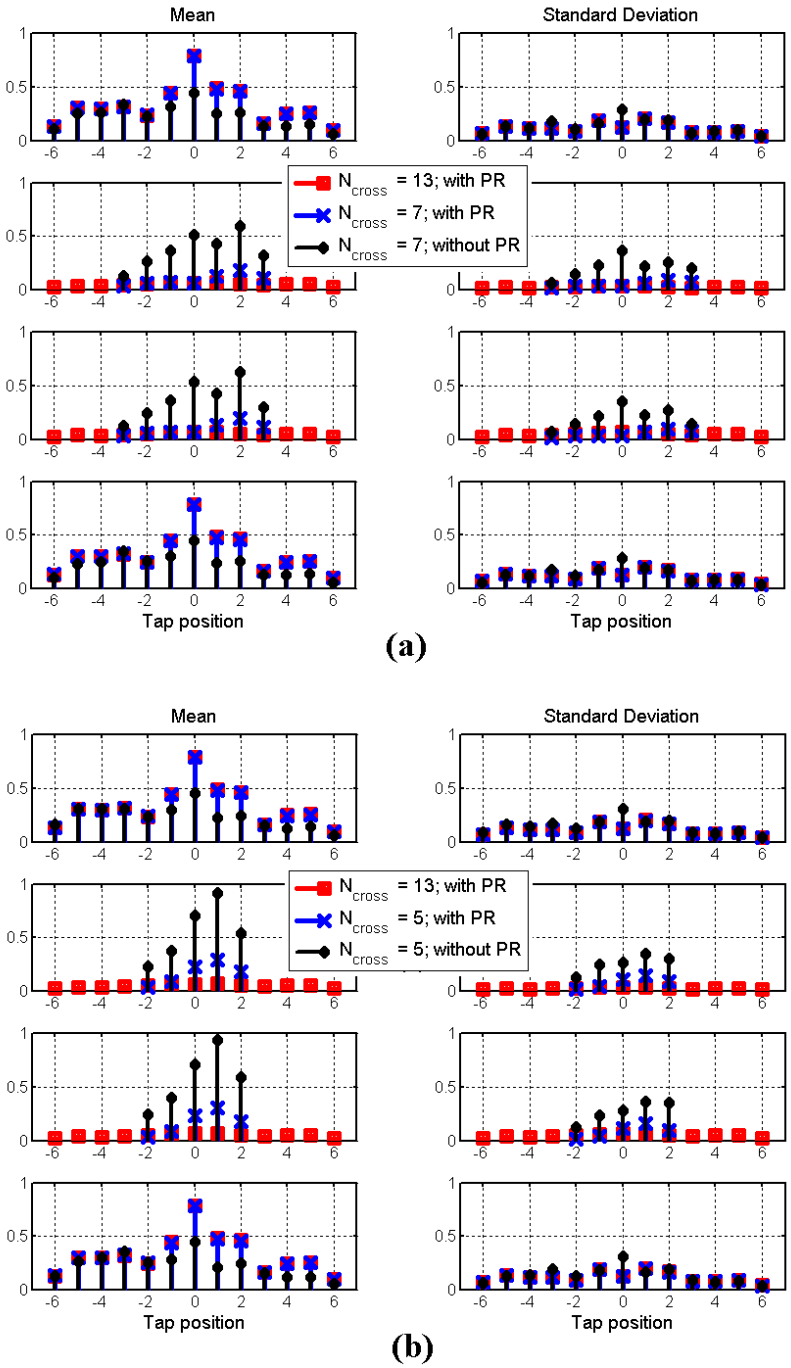

FIGURE 6.7 – The mean and the standard deviation of taps of MIMO-FIRs corresponding to the results in Fig. 6.6 when using (a) Black : 7 cross taps without SOP prerotation ; with tap initialization. Blue : 7 cross taps with SOP prerotation. Red : 13 cross taps with SOP prerotation. (b) Black : 5 cross taps without SOP prerotation ; with tap initialization only. Blue : 5 cross taps with SOP prerotation. Red : 13 cross taps with SOP prerotation. All above were generated using 10 degree-resolution blind SOP search. All above were generated using 10 degree-resolution blind SOP search, averaged over 245 unrepeated SOPs with 10 realizations for each SOP. PR : Prerotation.

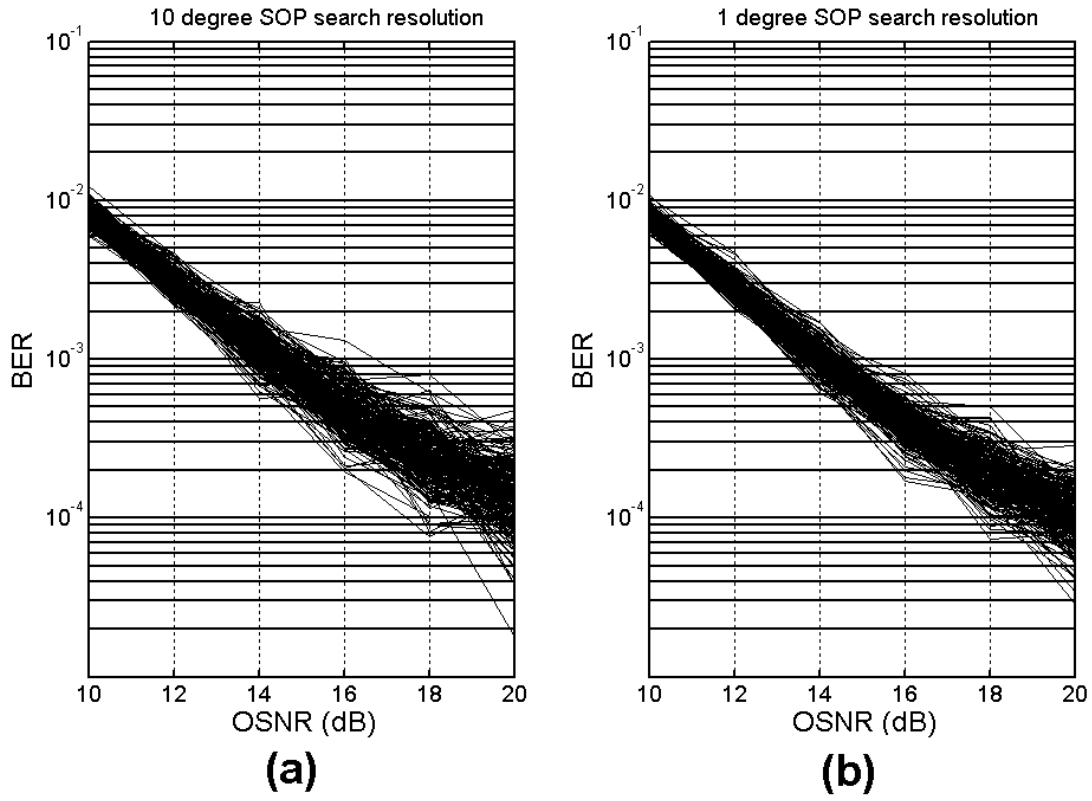

FIGURE 6.8 – Comparison in BER performance between (a) using 10-degree resolution blind SOP search and (b) 1-degree resolution blind SOP search for our reduced MIMO with 5 cross taps only. The above results were obtained after a 40-km FD-CDC.

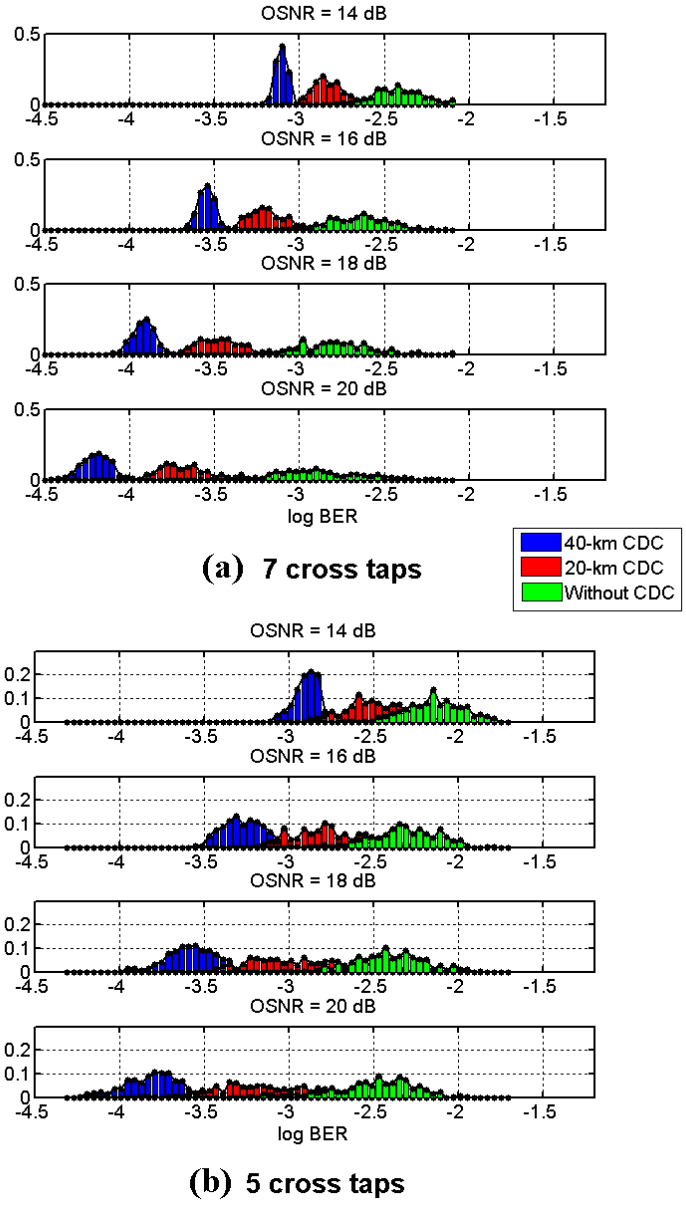

FIGURE 6.9 – The probability mass functions of BER conditioned on different OSNR values of our proposed DSP after 40 km SMF transmission, with 10-degree resolution in our coarse SOP estimation. Green : 5 cross taps with zero CDC (residual CD = 680ps/nm) ; Red : 5 cross taps with 20-km CDC (residual CD = 340ps/nm) ; blue : 5 cross taps with 40-km CDC ; BER curves correspond-ing to 245 different SOPs were shown for each case.

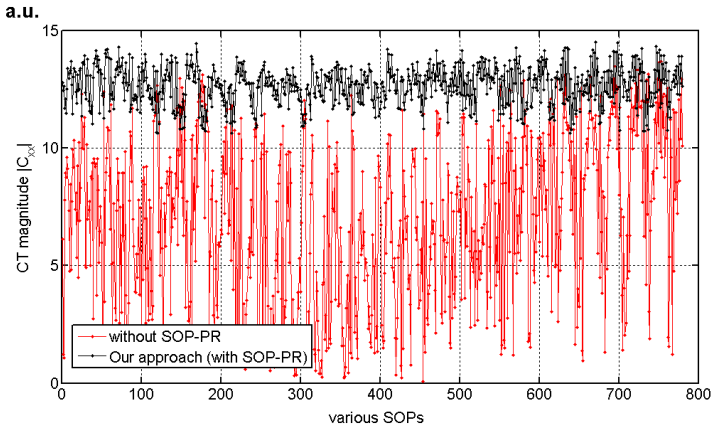

FIGURE 6.10 – Comparison of clock tone magnitudes between the DSP without (red) and with SOP-PR (black ; our proposed DSP) over 780 various unrepeated SOPs generated by a polarization synthesizer at OSNR of 10 dB. A.u. : arbitrary unit.

Chapitre 7

Extended Kalman Filter-based SOP Pre-rotation

In Chapter 6 we proposed a blind SOP search (parameterized by polarization orientation angle and phase delay between X and Y polarizations) to estimate the SOP of the received signal. The SOP was used for polarization pre-rotation to accelerate convergence and to reduce the hardware complexity of the conventional MIMO. Nevertheless, the blind SOP search requires memory to store the intermediate results calculated for each combination of parameters. This chapter is an extension of Chapter 6. Here, we propose a SOP tracking using an extended Kalman filter to reduce the hardware complexity introduced by the blind SOP search in Chapter 6.

The present chapter shares the same background and motivation with Chapter 6, which was given in Section 6.1. We will first specify our contributions. We will review the formulation of a conventional discrete-time extended Kalman filter. Then we will discuss the formulation of our proposed low-complexity Kalman filter based on minimization of s_1, as well as its implementation form. Finally, we will go through the experimental setup and our results for a polarization-scrambled (at 50 kHz) 32 Gbaud DP-QPSK system.

7.1 Contributions

SOP prerotation was first introduced for pure Stoke-space demultiplexing, primarily for applications in metrology [63], despite its high computational complexity in estimating Stokes space parameters for the input SOP. Appendix H provides a literature review for pure Stoke-space demultiplexing. Real-time implementation precludes high computation complexity and in Chapter 6, we thus proposed the use of a very coarse (with 10-degree resolution) blind search technique for SOP estimation. Nevertheless, blind search techniques are still considered computationally costly because the SOP estimation requires the averaging of resultants

from all SOP candidates, as shown in Fig. 6.3.

Kalman filter is well-known for its one-step operation, i.e., only the estimate(s) from the previous iteration is required to estimate or predict the parameters in the following iteration. In the present chapter, we propose to use a low-complexity discrete-time extended Kalman filter (EKF) minimizing resultant s_1 in order to track the inverse Jones matrix of an optical channel, reducing the memory depth required in our proposed blind SOP search in Chapter 6. Our proposed algorithm is updated only at every ASIC clock period and can be applied to real-time data.

The SOP prerotation allows us to obtain the clock information before MIMO, reducing the feedback delay to the ADCs. We show that this EKF allows us to keep maintaining a good clock tone magnitude level (which is important for timing phase estimation), has significant complexity reduction compared to the sample-wise multiplier blocks in timing-phase estimation algorithms [60, 62], and preserves all the advantages mentioned in Chapter 6, such as avoiding the singularity problem and long convergence of conventional MIMO, and reducing the subsequent MIMO-filter taps.

7.2 Equations of Extended Kalman Filter

In this section, equations for the extended Kalman filter will be revisited. We will adhere to the formulations from Szafraniec et al. [63], which uses the following notations :

- $\underline{x}_k$ denotes the true parameter (also called the state vector).
- $\underline{\hat{x}}_k$ is the posteriori estimate of $\underline{x}_k$, which is an estimated state vector after the adjustment using the knowledge of measurement (as a posterior information). It is generated in the **measurement update equation**. We use this vector for our estimated parameters.
- $\underline{\hat{x}}_k^-$ is the priori estimate of $\underline{x}_k$, which is an estimated state vector after the adjustment using the knowledge of parameter dynamics (as a prior information). It is generated in the **time update equation**. In our proposed DSP, $\underline{\hat{x}}_k$ is the same as $\underline{\hat{x}}_k^-$ as shown below.

7.2.1 Stochastic System

In the following, we start the standard formulation with scalar quantities for simplicity. The formulation holds also for vector quantities.

We assume that our true time-varying parameter x_k is described by a linear time difference equation as follow :

$$x_k = Ax_{k-1} + Bu_{k-1} + \zeta_{k-1}, \tag{7.1}$$

where the current state x_k evolves from the previous state x_{k-1} through matrix A, u_k represents a perturbation that alters the parameters through matrix B, and ζ_{k-1} represents the noise in the system. Equation (7.1) is called the **process equation**, and ζ_{k-1} is called the process noise.

Usually, one may not be able to observe the time-varying parameter directly upon measurement, i.e., x_k is hidden in the measurement system h. We can only access the measurement output :

$$z_k = h(x_k) + \xi_k \tag{7.2}$$

where ξ_k is the measurement noise. Function h can be either a linear or a nonlinear function. In this chapter, we only consider a nonlinear function for h. Equation (7.2) is called the **measurement equation**. The following assumptions are used :

1. ζ_k is an uncorrelated sequence of zero mean random vectors with process-noise covariance variance Q.
2. ξ_k is an uncorrelated sequence of zero mean random vectors with measurement-noise covariance variance R.
3. The initial state x_0 has mean m_0 and covariance P_o
4. ζ_k, ξ_k and x_0 are mutually uncorrelated.

The knowledge of matrices A, B and h (prior information) is crucial for us to apply the extended Kalman filter for tracking parameters.

7.2.2 Extended Kalman Filter Algorithm

Based on the above stochastic system, one can apply the standard extended Kalman filter step-by-step. In the following, we will only revisit the procedures of EKF. For the EKF derivation, please refer to the standard textbook [47].

The EKF algorithm estimates the state vector (the desired parameters) using the time update equations and the measurement update equations. The **time update equations** for state vector and for error covariance are :

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}, \tag{7.3}$$

$$P_k^- = AP_{k-1}A^{*T} + Q, \tag{7.4}$$

respectively, where P_{k-1} is the covariance matrix of estimation error, P_k^- is the covariance matrix of estimation error after the adjustment using the knowledge of parameter dynamics A and process noise Q. The **measurement update equations** are :

$$K_k = P_k^- H^{*T}(HP_k^- H^{*T} + R)^{-1}, \tag{7.5}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-)), \tag{7.6}$$

$$P_k = P_k^- - K_k H P_k^-. \tag{7.7}$$

where H is the Jacobian matrix of partial derivatives of function h with respect to state variables of x_k. K_k in (7.5) refers to Kalman gain. The term, $z_k - h(x_k^-)$, is called the innovation, or the instantaneous estimation error. The sequence of EKF is as follows :

1. **Initialization at** $k = 1$: Substitute x_0 into (7.3) and P_0 and Q into (7.4).
2. **Kalman gain update** : Calculate the Kalman gain in in (7.5) using the result in (7.4) and R.
3. **(a) Update estimate** $\hat{x}_k$: Substitute $\hat{x}_k^-$ of (7.3), the new measurement output z_k, Kalman gain K_k calculated in step 2 into (7.6).

   (b) Update State vector : In parallel to (a), Substitute P_k^- of (7.4) and Kalman gain K_k calculated step 2 into (7.7).
4. Repeat the above for the next time index k.

7.3 EKF-based SOP Pre-rotation

Fig. 7.1a shows our proposed DSP flow, SOP pre-rotation (PR) is performed by multiplying the CD-compensated signals with a 2×2 complex matrix which represents the inverse Jones matrix of the optical channel as shown in Fig. 7.1b. We parameterize this inverse Jones matrix, J_{PR}, using real numbers a, b, c and d [1].

$$J_{PR} = \begin{bmatrix} a + jb & c + jd \\ -c + jd & a - jb \end{bmatrix} \tag{7.8}$$

Let X_{in} and Y_{in} be the complex signals after CD compensation for polarizations X and Y. After SOP-PR using the Jones matrix, J_{PR}, the output signals are

$$\begin{bmatrix} X_{out} \\ Y_{out} \end{bmatrix} = \begin{bmatrix} a + jb & c + jd \\ -c + jd & a - jb \end{bmatrix} \begin{bmatrix} X_{in} \\ Y_{in} \end{bmatrix} \tag{7.9}$$

7.3.1 Our stochastic system

We start our EKF derivation based on the work of Szafraniec et al. [63] revisited in the previous section. Our aim is to track state vector $x_k = [\text{a b c d}]^T \in \mathbb{R}^{4\times 1}$. **Different from the conventional approach**, we set $A = I_4$, where I_N is an identity matrix of N dimensions,

1. A 2×2 complex matrix should be parameterized by eight real numbers; a unitary matrix requires only four real numbers.

and assume the absence of external perturbation, i.e., $u_{k-1} = 0$. Our new stochastic system, corresponding to (7.1) and (7.2), becomes

$$\underline{x}_k = A\underline{x}_{k-1} + \underline{\zeta}_{k-1}, \tag{7.10}$$

$$\underline{z}_k = \underline{h}(\underline{x}_k) + \underline{\xi}_k, \tag{7.11}$$

with $\underline{\zeta}_k \in \mathbb{R}^{4\times1}$, $\underline{\xi}_k \in \mathbb{R}^{2\times1}$, $\underline{h}(\underline{x}_k) \in \mathbb{R}^{2\times1}$ and $\underline{z}_k \in \mathbb{R}^{2\times1}$. We set two constraints on our nonlinear function h to force our J_{PR} to the inverse Jones matrix.

1. **Minimization of** s_1 : As with our blind SOP estimation in Chapter 6, J_{PR} pre-rotate the SOP of the input signals after CDC such that s_1 of the output signals from J_{PR} is close to zero :
$$s_1 =< |X_{out}|^2 - |Y_{out}|^2 >= 0 \tag{7.12}$$
2. **Unitary matrix** J_{PR} : assuming a moderate polarization dependent loss in optical channel, the determinant of J_{PR} should approach unity, such that
$$\left|J_{PR}\right| = a^2 + b^2 + c^2 + d^2 = 1. \tag{7.13}$$

From these two constraints, we define the nonlinear function h in the measurement equation as

$$h(\underline{x}_k) = \begin{bmatrix} < |X_{out}|^2 - |Y_{out}|^2 > \\ a^2 + b^2 + c^2 + d^2 \end{bmatrix} \in \mathbb{R}^{2\times1}, \tag{7.14}$$

where the first element of h is a nonlinear function of the output signals of the inverse Jones matrix, while the second element is a nonlinear function of the state variables.

Then, we force our state variables to be the entries of the desired Jones matrix by setting the measurement output as :

$$\underline{z}_k = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{R}^{2\times1}. \tag{7.15}$$

Different from the conventional approach, our measurement outputs are not from the physical measurement, but two constants that we want to obtain. Note that we talk about the stochastic system here only, and therefore the variable a, b, c and d are the true parameters.

7.3.2 Our EKF Algorithm

The time update equation for state x_k in (7.3) becomes

$$\underline{\hat{x}}_k^- = \underline{\hat{x}}_{k-1}, \tag{7.16}$$

where $\underline{\hat{x}}_k^- = \begin{bmatrix} \hat{a}_k & \hat{b}_k & \hat{c}_k & \hat{d}_k \end{bmatrix}^T \in \mathbb{R}^{4\times1}$. The time update equation for error covariance matrix in (7.4) becomes

$$P_k^- = P_{k-1} + Q. \tag{7.17}$$

The measurement update equations in (7.5), (7.6) and (7.7) become, respectively,

$$K_k = P_k^- H^{*T}(HP_k^- H^{*T} + R)^{-1}, \tag{7.18}$$

$$\underline{\hat{x}}_k = \underline{\hat{x}}_k^- + K_k \begin{bmatrix} < |X_{out}|^2 - |Y_{out}|^2 > \\ \hat{a}_k^2 + \hat{b}_k^2 + \hat{c}_k^2 + \hat{d}_k^2 - 1 \end{bmatrix}, \tag{7.19}$$

$$P_k = P_k^- - K_k H P_k^-, \tag{7.20}$$

with $\underline{\hat{x}}_k^- \in \mathbb{R}^{4\times 1}$, $\underline{\hat{x}}_k \in \mathbb{R}^{4\times 1}$, $P_k^- \in \mathbb{R}^{4\times 4}$, $P_k \in \mathbb{R}^{4\times 4}$, $K_k \in \mathbb{R}^{4\times 2}$, $Q \in \mathbb{R}^{4\times 4}$, $R \in \mathbb{R}^{2\times 2}$. The process-noise covariance variance Q and the measurement-noise covariance variance R determine the tracking speed and the convergence speed of the algorithm. We will adjust their values in our experimental DSP as shown later in this chapter. The Jacobian matrix H can be calculated using (7.13) :

$$H(\underline{\hat{x}}_k^-, X_{in}, Y_{in}) = \begin{bmatrix} f(\hat{a}_k, \hat{c}_k, \hat{d}_k) & f(\hat{b}_k, \hat{d}_k, -\hat{c}_k) & f(-\hat{c}_k, \hat{a}_k, -\hat{b}_k) & f(-\hat{d}_k, \hat{b}_k, \hat{a}_k) \\ 2\hat{a}_k & 2\hat{b}_k & 2\hat{c}_k & 2\hat{d}_k \end{bmatrix} \in \mathbb{R}^{2\times 4} \tag{7.21}$$

where

$$f(p,q,r) \triangleq 2p|X_{in}|^2 - 2p|Y_{in}|^2 + 4Re[(q+jr)X_{in}^* Y_{in}] \in \mathbb{R}^{1\times 1} \tag{7.22}$$

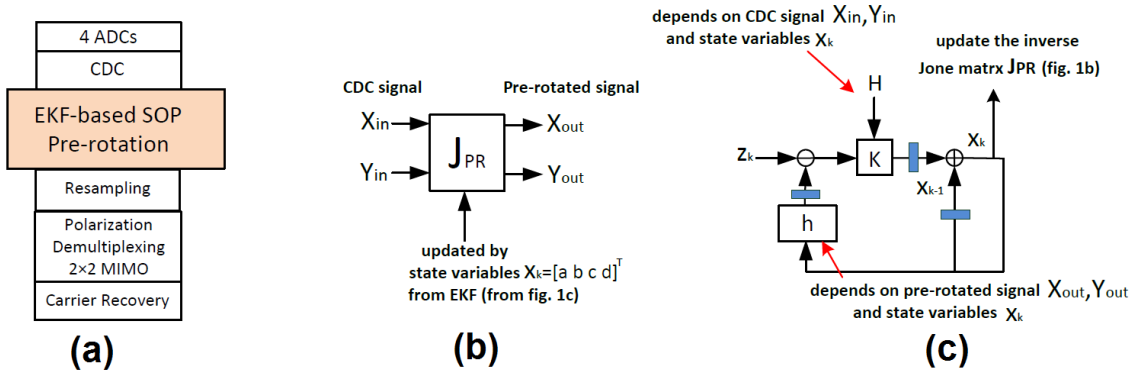

FIGURE 7.1 – (a) Our proposed DSP flow. ADC : analog-to-digital converter, CDC : chromatic dispersion compensation, SOP : state-of-polarization. EKF : extended Kalman filter (b) SOP prerotation (c) our extended Kalman filter flow chat. Blue rectangles : pipelining delays.

7.4 Implementation

As SOP rotation is a slowly varying effect compared to the symbol rate, the state variables are updated and are used to renew J_{PR} only at every ASIC clock period (e.g. hundreds of symbol

durations), compared to the sample-wise multiplier blocks in [60, 62]. No parallelization is required. The squaring of input power and output power for equations (7.19) and (7.22) are the only sample-by-sample operations, which can be implemented easily using look-up tables (LUT). The first row element of the last term in (7.19) is generated by a block average over an ASIC clock. For pipelining, the extended Kalman filter is equivalent to the following loop equation with feedback delay

$$\underline{x}_k = \underline{x}_{k-D} + K_{k-3D}\left[\begin{pmatrix}0 & 1\end{pmatrix}^T - \underline{h}(\underline{x}_{k-D}, X_{out}, Y_{out})\right]^T \tag{7.23}$$

where D (the number of delayed samples) is 50, corresponding to an ASIC rate of 800 MHz for a sampling rate of 40 Gsa/s. From now on, we drop the hat sign and the superscript for simplicity. Fig. 7.2 shows the implementation of our EKF-based SOP-PR. The increments of parameters (the second term of RHS of (7.23)) are delayed by $3D$. The estimated state variables refresh J_{PR} in (7.14) to pre-rotate data before MIMO processing. The pre-rotated signal is used to generate an innovation per (7.19). The innovation is multiplied with the Kalman gain K_{k-3D} . The Kalman gain is delayed by $3D$ because the estimated state variables go through the square-law operation to generate input signal powers. Then, look-up tables (LUTs) are used to find the Kalman gain, followed by the multiplication of Kalman gain with the innovation.

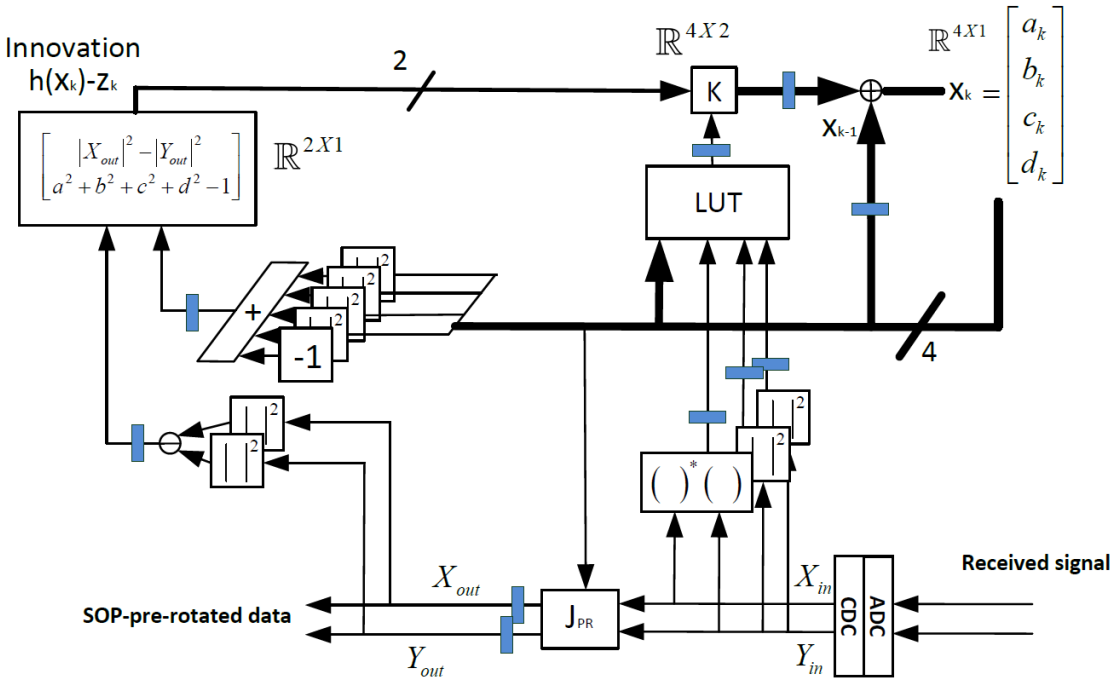

FIGURE 7.2 – Implementation diagram of our EKF-based SOP-PR

7.5 Performance Analysis

7.5.1 Experimental Setup

The performance of our proposed DSP is examined using the experimental setup shown in Fig. 6.5. A 32 Gbaud DPQPSK signal was generated by a polarization multiplexing emulator (dashed, orange), with laser sources of 10 kHz linewidth. A polarization synthesizer (Agilent N7786B) was used to rotate the SOP of the transmitted signal at 50 kHz. After passing through an erbium doped fiber amplifier (EDFA) and a 40 km standard single-mode fiber (SMF), the received signal was coherently detected by an integrated coherent receiver having a bandwidth of 22 GHz. The signal was captured by a sampling scope at 40 GSa/s with limited (16 GHz) bandwidth.

7.5.2 Offline DSP

For offline DSP, we followed the DSP flow in Fig. 7.1 to perform CD compensation first. We applied our proposed EKF-based SOP-PR (operated at an ASIC rate of 800 MHz) directly before resampling. At this point, we looked into the results of the CT tracking performance, as explained in the next paragraph. To see the application of SOP pre-rotation, we performed retiming and resampling to 2 samples per symbol, followed by polarization demultiplexing (using our proposed reduced-complexity MIMO in Chapter 6), phase recovery, detection and bit error counting.

In order to compare the clock tone magnitude with and without KF-based SOP-PR (importance of SOP-PR), we resampled the signal to 2 samples per symbol and extracted the clock tone using frequency-domain (FD) Godard timing phase estimation (TPE) over 1000 samples [57]. Please refer to Appendix G for Godard TPE. The results of the CT tracking performance are shown in Fig. 7.3, while those of the BER performance using our previously propsoed reduced-complexity MIMO are shown in Fig. 7.4.

7.5.3 Experimental Results

Our experimental results were generated based on a polarization-scrambled (at a speed of 50 kHz) 32 GBaud DP-QPSK signal. The covariance matrices for process noise Q and for measurement noise R are set via trial and error to 0.5 I_4 and 5000 I_2, respectively. The physical meanings of these two parameters are related to the convergence speed and the tracking bandwidth of the algorithm, which was well explained in [63]. However, for large signal distortion (refer to our case where 32 GBaud QPSK was detected by a 16 GHz receiver), we had to resort to a trial and error on our measurement. We chose values for these two parameters to achieve a stable level of clock tone magnitude with short convergence time, as shown in Fig. 7.3b. The initial state is set to $[1\ 0\ 0\ 0\ ]^T$.

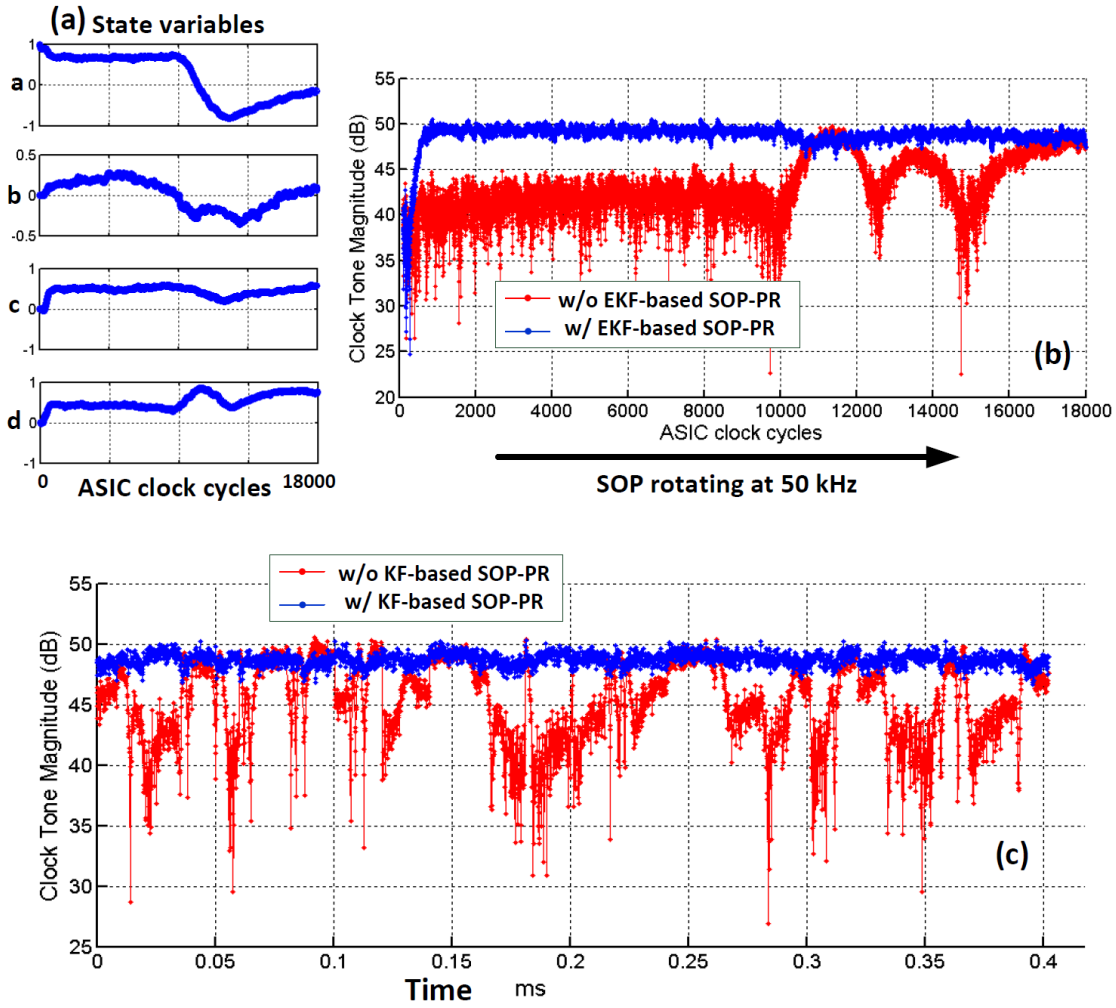

FIGURE 7.3 – (a) Evolution of state variables (parameters of Jones matrix). (b) Comparison of clock tone magnitudes within 18000 ASIC clock cycles before and after EKF-based SOP-PR. (c) Comparison of clock tone magnitudes over a longer duration (over 332000 ASIC clock cycles) before and after EKF-based SOP-PR. All plots were generated using polarization-scrambled (at 50 kHz) 32 Gbaud DP-QPSK at OSNR = 16 dB.

Fig. 7.3a shows the evolution of the estimated state parameters $\hat{a}_k$, $\hat{b}_k$, $\hat{c}_k$ and $\hat{d}_k$ within 18,000 ASIC clock cycles (ASIC rate = 800 MHz). Each point of the curves corresponds to the estimated values from our proposed EKF algorithm at every 50 samples (recall that the received signal was sampled at 40 Gsa/s , while the EKF algorithm operates at 800 MHz only, meaning that the algorithm was applied every $\frac{40\text{G}}{800\text{M}} = 50$ samples.)

Corresponding to Fig. 7.3a , Fig. 7.3b shows the CT magnitudes before (red) and after (blue) KF-based SOP-PR. The CT magnitude fluctuates with SOP rotation (red, i.e., without EKF-

based SOP-PR). During the acquisition stage (the first 1000 ASIC clocks), the two curves (red and blue) overlap. Once EKF adapts to a correct SOP, our proposed EKF-based SOP-PR restores the CT magnitude and maintains the CT magnitude at a high level. This helps reduce jitter and eliminate system failure of coherent receivers. To shorten initial convergence time, we can simply use blind SOP search proposed in Chapter 6.3 to initialize the state variables (not shown here). Fig. 7.3c is similar to Fig. 7.3b, but showing the clock-tone magnitudes before and after EKF-based SOP-PR over a longer duration (over 332000 ASIC clock cycles, or $332000 \times \frac{1}{800\text{MHz}} = 0.415$ ms).

To show the advantage (application) of our proposed EKF-based SOP-PR, Fig. 7.4 shows the BERs versus time of our proposed reduced-complexity MIMO filter taps (with $N_{cross} = 7$, proposed in Chapter 6) without (red) and with (blue) EKF-based SOR-PR. To verify that our proposed EKF can pre-rotate SOP correctly, the evolution of the BER values was shown at a fixed OSNR of 16 dB, measured every minute, taken over 120 minutes, i.e. each curve consists of 120 points. At each point, 80000 DP-QPSK symbols were used to calculate a BER value. Our proposed technique results in a lower BER (below 1e-3), because the EKF tracks the SOP rotation, and J_{PR} coarsely reject the SOP change. Therefore, the cross taps of MIMO can be significantly reduced.

Fig. 7.5 shows the averaged BERs (over 120 minutes) versus OSNRs of reduced-complexity MIMO filter taps without (red) and with (blue) EKF-based SOR-PR, compared with the best performance given by a full-complexity MIMO (black). The BER value at each OSNR was calculated by using all 120 captures within 120 minutes, i.e., $120 \times 80000 = 9.6$ million DP-QPSK symbols. Our proposed technique improves the OSNR penalty at BER of 1e-3 by 1.5 dB for reduced-complexity MIMO filter, but degrades by 0.5 dB compared to a full-complexity MIMO. This extra link budget is counter-balanced by the reduced complexity of our proposed algorithm, and is still within the working OSNR range in short-reach scenarios.

7.6 Conclusion

In this chapter, we have extended our work in Chapter 6 to use SOP pre-rotation to solve the problems of the conventional MIMO. The difference in this chapter is that we have proposed a low complexity discrete-time extended Kalman filter operated at ASIC rates to track the Jones matrix of the inverse of an optical channel (the polarization effects) based on minimization of s_1. Compared to the blind SOP search proposed in Chapter 6, the extended Kalman filter can save memory and redundant computations.

In Chapter 6, an experimental demonstration was performed for various static SOPs, followed by a complexity-performance trade-off to emphasize the importance of SOP-PR for our proposed reduced-complexity MIMO for short-reach scenarios. This "proof-of-concept" showed that the weakened clock tone can also be restored at the same time for various SOPs,

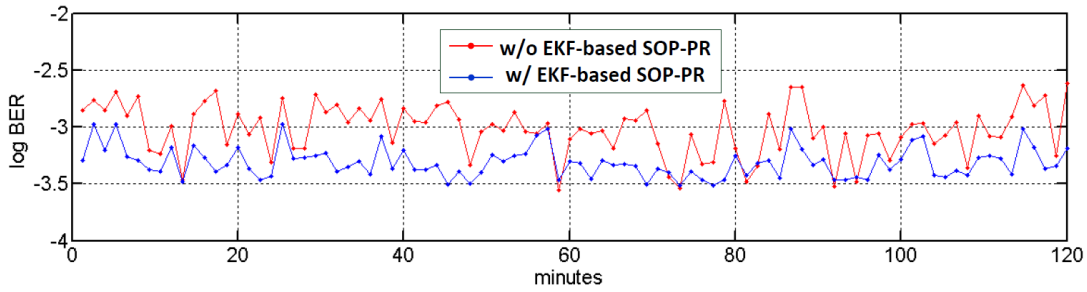

FIGURE 7.4 – Comparison of bit error rates (measured every minute) of using reduced-complexity MIMO filter taps ($N_{cross} = 7$) proposed in Chapter 6 with and without EKF-based SOP-PR. All plots were generated using polarization-scrambled (at 50 kHz) 32 Gbaud DP-QPSK at OSNR = 16 dB.

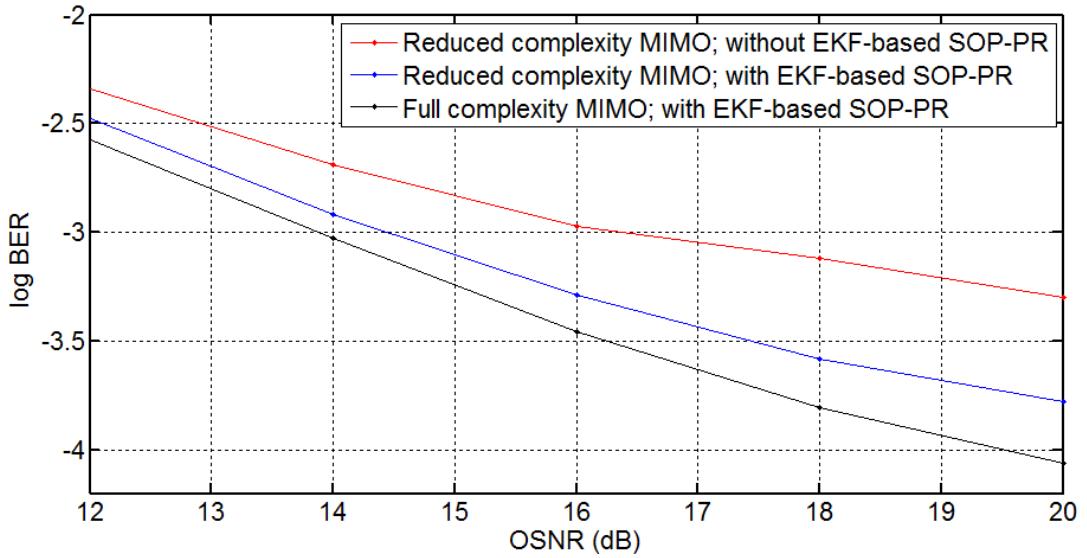

FIGURE 7.5 – Comparison of bit error rates versus OSNRs of using reduced-complexity MIMO filter taps ($N_{cross} = 7$) in Chapter 6 with and without EKF-based SOP-PR, and using a full-complexity MIMO with EKF-based SOP-PR. All points were generated using polarization-scrambled (at 50 kHz) 32 Gbaud DP-QPSK

suitable for applications for burst-mode receivers in short-reach scenarios. In this chapter, our experimental demonstration focused on the performance of our proposed DSP in clock tone recovery. We tested our proposed SOP tracking and previously proposed reduced-complexity MIMO with dynamic SOPs (scrambled at 50 kHz).

Our experimental results of a polarization-scrambled 32 Gbaud DP-QPSK system show that

the clock tone magnitude can be restored by our proposed extended Kalman-filter-based SOP pre-rotation at a 800 MHz ASIC rate. This avoids the clock tone loss due to the polarization effect of the optical channel, and enhances the bit-error rate performance of using a reduced-complexity MIMO for digital polarization demultiplexing.

Chapitre 8

Conclusion

In this thesis our first contribution was experimentally showing, for the first time, that TeraXion's narrow-linewidth lasers gives BER improvement in real-time phase recovery (parallel and pipelined decision-directed maximum likelihood estimation) that consists of a feedback loop. Using a fiber Bragg grating suppresses the frequency noise spectral level of narrow-linewidth semiconductor lasers such as external cavity lasers or integrated tunable laser assemblies. The laser source permits greater parallelization, e.g., increase from 16 to 20, to reduce the hardware processing rate from 312.5 to 250 MHz. We pointed out that the use of laser linewidth to predict system performance is no longer appropriate to quantify laser phase noise for a laser with frequency noise suppression. We suggested the use of frequency noise power spectral density levels to quantify the impact of laser phase noise for real-time phase tracking with feedback delay. Offline experimental demonstrations often neglect feedback delay.

Hardware parallelization and pipelining are required to implement real-time systems, but impose delay on the feedback path of the real-time decision-directed phase recovery, reducing its tracking bandwidth. For current electronics (i.e. ASIC rates between 200 MHz and 400 MHz are used because of optimal power consumption [29]), we found that the laser frequency noise between 10 MHz and 100 MHz affects the BER performance of real-time phase tracking for a single-polarization, single-carrier 5 Gbaud 64-QAM system. We expect that this conclusion also holds for higher symbol rates as well, as a higher symbol rate requires higher parallelization in the current CMOS technology.

Our second contribution was pointing out, for the first time, that the standard frequency offset compensation algorithm could cause real-time phase tracking failure for 64-QAM in the presence of sine tones in laser sources below 1 MHz. We demonstrated experimentally the impact of sinusoidal laser phase noise on phase recovery in the presence of parallelization and pipelining delay. Our demonstration applied parallel and pipelined DD-MLE to a single-polarization, single-carrier 5-Gbaud 64-QAM system, taking into account the popu-

lar NDA-FFT for frequency-offset compensation and static equalization. We experimentally investigated, for the first time, the ranges of the FM amplitude and FM frequency of laser sources to avoid real-time phase tracking failure. Together with our simulation results, we found that the frequency modulated (FM) amplitude and FM frequency of laser sources should be smaller than 1.5 MHz and 75 kHz, respectively, to avoid tracking failure.

Our third contribution was proposing a novel DSP technique to solve the problems of conventional MIMOs for polarization demultiplexing and reducing the power consumption in short-reach coherent communications systems, where DGD is present but not prominent. We proposed a very coarse SOP pre-rotation before MIMO using an inverse Jones matrix based on the minimization of only a single Stokes parameter (s_1) via a blind SOP search. This SOP pre-rotation can coarsely reject the polarization crosstalk before the subsequent polarization demultiplexing DSP. It brings advantages compared to the conventional MIMO-FIR approaches : only a single Stokes parameter instead of all three Stokes parameters, greatly reducing the computational effort for estimating SOP compared to other Stokes-space approaches. It avoids the problem of long MIMO convergence and singularity for certain SOPs, which makes the DSP very suitable to burst-mode receivers normally in packet-switching based short-reach or metro networks. As the SOP pre-rotation reduces polarization coupling, we are then at liberty to reduce the number of MIMO cross-taps (the off-diagonal components in a Jones matrix), leading to a significant reduction in number of MIMO-FIR complex multipliers. The SOP pre-rotation also brings the benefit of restoring clock tones for timing phase estimation (TPE) even before MIMO. The hardware overhead for coarse SOP estimation is easily counter balanced by savings in MIMO complexity. We have also experimentally presented a tradeoff between hardware reduction and performance degradation in the presence of residual chromatic dispersion for short-reach applications.

Our fourth contribution was proposing a low-complexity discrete-time extended Kalman filter (EKF) minimizing resultant s_1 in order to track depth required in our proposed blind SOP search in our third contribution. Our proposed algorithm is updated only at every ASIC clock period and can be applied to real-time data. We experimentally showed that this EKF allows us to keep maintaining a good clock tone magnitude level (which is important for timing phase estimation), has significant complexity reduction compared to the existing sample-wise multiplier blocks in timing-phase estimation algorithms, and preserves all the advantages mentioned in our third contribution, such as avoiding the singularity problem and long convergence of conventional MIMO, and helps reduce the subsequent MIMO-filter taps.

8.1 Future Work

Coherent systems usually have higher performance than other systems, but power consumption is always a concern for implementing DSP on ASIC chips, especially for metro networks

and short-reach applications where the device density (device per user) is high. Crivelli et al. [5] summarized, for the first time, the power consumption of different DSP stages in a long-haul DP-QPSK coherent chip : the highest power consumption of a chip is due to the bulk CD compensation consumes (11 W), while the second highest power consumption is due to the equalization using MIMO-FIR (6.95 W). Undoubtedly, MIMO-FIR becomes a major consumption in short-reach systems where CD is small, and our DSP designs in Chapters 6 and 7 are to further reduce the power consumption of digital coherent receivers for short-reach applications. However, the reduction of power consumption is only the calculation of the number of complex multiplications used in our new DSP, which is a crude evaluation only. The true power consumption varies with implementation. One of our future work is to examine the performance of our DSP design in Chapters 6 and 7 via implementation in FPGA experimentally before the design is realised in ASIC for further circuit optimization for power saving.

It is well-known that the convergence time is determined mainly by MIMO-FIRs in burst-mode receivers for packet-switching applications in short reach systems [33]. Reducing convergence time helps enhancing the chip power efficiency for useful packet content. We may calculate the acquisition times required by our proposed DSP in Chapters 6 and 7 for different SOPs so that the duration of headers can be tailored made for each data packet.

Dependence on modulation format is also a concern of designing DSP. Our proposed DSP for 32 Gbaud DP-QPSK (100 Gbps system) relies on the signal representation of dual-polarization signals in Stokes space, forming a lens object on the Poincaré sphere. Higher-order modulation formats are more susceptible to receivers' electrical bandwidth, resulting in a distorted lens object causing inaccurate SOP estimate. One may suggest to employ 16 Gbaud DP-16-QAM to achieve 100 Gbps systems, as the reduced symbol rate helps saving parallelization and thus power consumption. One of our future work is to make our SOP estimate generic to 16-QAM or higher.

Annexe A

Phase Estimation and Phase Tracking

In this appendix, we discuss the definition of phase estimation and phase tracking in a stricter sense. The unknown quantity or variable of a system that one desires to find out is called a parameter. This parameter may be random in nature, and therefore statistical estimation methods are required. For example, in Chapter 2 and 3, the parameter of concern is the laser phase noise.

A.1 Theoretical viewpoint

There are two kinds of knowledge about a random process to understand its nature. Firstly, the **ensemble statistics** that give the probability distribution of the random process at a certain time instant, i.e., a random variable only. Examples are mean and variance of a distribution.

Secondly, as a random experiment proceeds with time, i.e., the same experiment is repeated, the randomness may be the same or different at different time instants [1]. This refers to a random process. Usually, we are interested in the relationship between the repeated experiments (i.e., between the random variables at different time instants), as we aim at describing the randomness in the past or predicting what happens in future. The repeated experiments at different time points can only be either independent to, or correlated with each other, which correspond to independent random processes, and Markov processes, respectively. Our effort lies on finding out the **process properties** (in a continuous sense) or the **time series properties** (in a discrete sense) which gives the temporal dynamics or system dynamics that governs how a random process behaves, i.e., the correlation between the repeated expe-

1. For example, the stock market behaves differently in times slots from 9 and 10 am and from 10 to 11 am.

riments at different time instants. For linear systems, the temporal dynamics can usually be represented by linear differential equations.

Strictly speaking, one has to find out all the higher-order joint moments of the random processes at different time points, in case that the random process shows time-varying statistics, i.e., non-stationary random process. In particular, for a second-order stationary (or called wide-sense stationary) random process, one can characterize a random process using stochastic tools such as autocorrelation function and power spectral density to capture its properties based on its ensemble statistics and time-series properties.

In mathematics, we define random variables to be estimated as parameters if their statistical properties is time-invariant. Time-varying parameters are usually called signals [47], which are random processes. **Parameter estimation** only requires the knowledge of the ensemble statistics of a random parameter to be estimated. In Chapter 2, one assumes that the parameter, i.e., laser phase noise, is approximately a constant over a temporal observation window, and **phase estimation** is performed using averaging to obtain the mean of the phase variation over a certain time. **Signal tracking** requires the knowledge of both the ensemble statistics and the temporal dynamics (time-series properties) of a random process to be estimated. In fact, in Chapter 2, if one wishes to have an optimal estimation of the signal, i.e., time-varying laser phase noise (following Brownian motion), **phase tracking** should be performed by making use of the properties of the Wiener process. Wiener filter and Kalman filter are two typical examples used for optimal signal tracking [47].

For example, the conventional Viterbi-Viterbi algorithm [2] requiring only a moving average belongs to the class of phase estimation because phase noise dynamics are not required. On the other hand, Viterbi-Viterbi algorithm followed by a Wiener filter [23] belongs to phase tracking because the phase noise dynamics are assumed to be a first-order difference equation (random walk). Intuitively, as phase tracking requires more knowledge about a random process to be estimated, phase tracking can give a more "optimal" phase estimate than phase estimation.

A.2 Commercial Systems

The techonology of optical coherent communications has been widely deployed, and publications from both academia and industries are becoming more industrial-oriented. For simplicity, phase estimation usually refers to the process of estimating a time-varying phase via moving average in a feedforward structure (without feedback). Phase tracking usually refers to the process of estimating phase with feedback, using the previous estimated phase to update or assist the current phase estimation.

2. Recall that Viterbi-Viterbi algorithm removes the data information by taking M^{th} power of the received signal, and therefore only phase is left as an unknown parameter. Please refer to Section 2.2.2.

Annexe B

Laser Phase Noise

In this appendix, we will first give the mathematical definitions of frequency noise and Wiener phase noise. We will see that Wiener phase noise is the integration of white Gaussian frequency noise. We will define the phase noise increment as well as power spectral density of the optical field. Phase noise increment is important for evaluating field power spectral density. In Section F.1, we make use of phase noise variance or field power spectral density for estimating laser linewidths. The most important equation in this appendix is B.14, especially for Chapter 2 and 3. We will see that the discretized Wiener process is described by a Gaussian random walk process. Finally, the procedures for generating Wiener phase noise in Matlab will be covered.

Definition of Frequency Noise

First, we define $\Delta f(t)$ as frequency noise of a laser source, an independent and identically distributed (i.i.d) zero-mean Gaussian random process with a variance of $\sigma^2_{\Delta f}$. The random process $\Delta f(t)$ has an autocorrelation function

$$R_{\Delta f}(r,s) = E\left[\Delta f\left(r\right)\Delta f\left(s\right)\right] = \sigma^2_{\Delta f}\delta\left(r-s\right), \tag{B.1}$$

where r and s are time variables, $E\left[\cdot\right]$ refers to the ensemble average (expectation) with respect to the joint distribution of $\Delta f\left(r\right)$ and $\Delta f\left(s\right)$, $\delta\left(r-s\right)$ is the Dirac delta function. When $r=s$, (B.1) becomes

$$R_{\Delta f}(r,r) = E\left[\Delta f\left(r\right)\Delta f\left(r\right)\right] = \sigma^2_{\Delta f}, \tag{B.2}$$

which is the variance of the random variable Δf.[1] The power spectral density (PSD) $S_{\Delta f}(f)$ of the random process Δf is

$$S_{\Delta f}(f) = \sigma^2_{\Delta f}, \forall f. \tag{B.4}$$

Since the PSD is flat, the corresponding Gaussian random process is called a white Gaussian process.

Definition of Wiener Phase Noise

The frequency of a signal is the derivative of the phase, hence phase noise can be written as

$$\theta(t) = 2\pi \int_0^t \Delta f(s)\, ds. \tag{B.5}$$

This process is non-stationary. For instance, it has a variance $\sigma^2_{\Delta f} t$ that is a function of absolute time t (p. 517 in [14]). Since integration is a linear operation, phase noise is also a Gaussian process. Strictly speaking, one should not comment directly on the variance of Wiener phase noise without defining a reference time point (where Wiener process starts). The value $\sigma^2_{\Delta f} t$ refers to the variance of Wiener phase noise from $t = 0$ up to time t.

Definition of Phase Noise Increment

From (B.5), we define the phase noise increment as

$$\theta_\tau(t) \triangleq \theta(t+\tau) - \theta(t) = 2\pi \int_t^{t+|\tau|} \Delta f(s) ds. \tag{B.6}$$

The phase noise increment over the time interval τ is a zero-mean Gaussian random process[2] with variance :

$$\sigma^2_{\theta_\tau} = E\left[|\theta_\tau|^2\right] = E\left[\left\{2\pi \int_t^{t+|\tau|} \Delta f(s) ds\right\}^2\right] = (2\pi)^2 \int_t^{t+|\tau|} ds \int_t^{t+|\tau|} du \cdot \sigma^2_{\Delta f} \delta(s-u). \tag{B.7}$$

Since $\int_t^{t+|\tau|} \delta(s-u)\, du = 1$ for $t \le s \le t+|\tau|$ and zero for elsewhere, we have

$$\sigma^2_{\theta_\tau} = (2\pi)^2 \sigma^2_{\Delta f} \int_t^{t+|\tau|} ds = (2\pi)^2 \sigma^2_{\Delta f} |\tau|. \tag{B.8}$$

Therefore, the autocorrelation function of phase noise increment depends only on time interval τ and not absolute time t, implying that the random process is WSS, and the Gaussian process $\theta_\tau(t)$ is stationary. Thus, $\theta_\tau(t), \theta_\tau(t+\tau), \theta_\tau(t+2\tau)$ are i.i.d., i.e., Wiener process has stationary independent increments.

1. As the autocorrelation is only a function of the lag τ, we can write

$$R_{\Delta f}(\tau) = \sigma^2_{\Delta f} \delta(\tau), \tag{B.3}$$

meaning that the random process is wide-sense stationary (WSS). As $\Delta f(t)$ is a Gaussian random process, it is not only WSS but also stationary. As the process is WSS, we can apply Wiener-Khintchine theorem to find its power spectral density (PSD) by taking the Fourier transform of the autocorrelation with respect to τ.

2. $\theta_\tau(t)$ is a Gaussian random process because of the linearity of integration

Power Spectral Density of Optical Field

We define the optical field of a laser source with constant amplitude and phase noise $\theta(t)$ as

$$A(t) \triangleq e^{j\theta(t)}. \tag{B.9}$$

Its autocorrelation function between two samples at time t and time s is

$$R_A(t,s) = E[A(t)A^*(s)] = E\left[e^{j\theta(t)}e^{-j\theta(s)}\right], \tag{B.10}$$

where $E[\cdot]$ refers to the ensemble average with respect to the joint distribution of the random variables $\theta(t)$ and $\theta(s)$ at time instant t and s respectively. Let $t = t + |\tau|, s = t$, using (B.8) and the characteristic function for zero-mean Gaussian random variable, the autocorrelation becomes

$$R_A(t+|\tau|,t) = R_A(\tau) = e^{-2\pi^2\sigma_{\Delta f}^2|\tau|}, \tag{B.11}$$

which is a function of time interval τ only and therefore the optical field of the laser source is WSS. The exponential autocorrelation [14] results in a well-known Lorentzian PSD, i.e.

$$S_A(f) = \int_{-\infty}^{\infty} R_A(\tau)\, e^{-j2\pi f\tau} d\tau = \frac{\sigma_{\Delta f}^2}{\left(\pi\sigma_{\Delta f}^2\right)^2 + f^2}. \tag{B.12}$$

The linewidth $\Delta\nu$ of a laser source is defined as the 3-dB width (two-sided) of the Lorentzian field PSD [7] :

$$\Delta\nu = 2\pi\sigma_{\Delta f}^2 = 2\pi S_{\Delta f}(f). \tag{B.13}$$

Note that the above linewidth definition is only for laser sources having pure Wiener phase noise, or equivalently white frequency noise, satisfying (B.4). (B.12) is not valid for colored frequency-noise PSD as discussed in Chapter 3. For Wiener phase noise, the linewidth can be found directly by referring to the level of frequency-noise PSD, and the phase noise increment over time T is thus, by substituting (B.13) into (B.8),

$$\boxed{\sigma_{\theta_T}^2 = E[|\theta(t+T) - \theta(t)|^2] = 2\pi\Delta\nu T.} \tag{B.14}$$

Discretization of Wiener Phase Noise

To perform digital signal processing on measurement data or to conduct simulation, one needs to consider discretization random process rather than continuous process. Here we would clarify how we deal with the discretized phase noise.

Wiener process in (B.5) means simply a first-order Markov process that obeys Langevin equation with an indepedent Gaussian noise source. (B.5) can be discretized as a sum process :

$$\theta_k = 2\pi \sum_{m=0}^{k-1} f_m, \tag{B.15}$$

where k is the time index taking every sampling period T_s, f_k is the discretized frequency noise, $\theta_0 = 0$ defines the starting point, $\theta_k = \theta(kT_s)$, $f_k = f(kT_s)$. In fact, (B.15) is a Gaussian random walk :

$$\theta_k = \theta_{k-1} + 2\pi f_k T_s. \tag{B.16}$$

Ideally, frequency noise $\Delta f(t)$ corresponds a phase noise increment $\theta_\tau(t)$ in (B.6) when τ approaches zero, i.e. $\Delta f(t) = \lim_{\tau\to 0} \frac{\theta_\tau(t)}{2\pi\tau}$. In practice or in simulation, τ is limited by sampling duration $T_s = 1/BW_s$, where BW_s is the sampling rate or the simulation bandwidth. Therefore, the discretized frequency noise becomes $f_k = \frac{\theta_{T_s}(kT_s)}{2\pi T_s}$, and, using (B.8), has a variance of $\Sigma^2_{\Delta f}$, where $\Sigma^2_{\Delta f} = \frac{(2\pi)^2\sigma^2_{\Delta f}T_s}{(2\pi T_s)^2} = \frac{\sigma^2_{\Delta f}}{T_s}$. This shows that the discretized frequency noise requires a rescaling to guarantee a proper variance. In Matlab, the discretized frequency noise should be generated using the following procedures :

1. generate an i.i.d. zero-mean normal vector $\underline{U}_k$;
2. specify the value of laser linewidth $\Delta\nu$;
3. calculate the variance using (B.13), i.e. $\sigma^2_{\Delta f} = \frac{\Delta\nu}{2\pi}$
4. multiply the vector with $\Sigma_{\Delta f}$ to make sure that the overall variance of the vector is still $\sigma^2_{\Delta f}/T_s$. The resultant vector, the discretized frequency noise vector, is $\underline{f}_k = \sqrt{\frac{\Delta\nu}{2\pi T_s}}\underline{U}_k$;
5. using (B.16), the discretized Wiener phase noise can be generated via $\theta_k = \theta_{k-1} + \nu_k$, where $\underline{\nu}_k = \sqrt{2\pi\Delta\nu T_s}\underline{U}_k$.

Annexe C

Reduction in Power Consumption due to Parallelization and Pipelining

This section reviews the reduction in power consumption of CMOS circuits due to pipelining and parallelization [[42], Ch. 3 and 10], discussed in Chapter 2. First, the power consumption P_{CMOS} of a CMOS (complementary mental-oxide-semiconductor) circuit can be approximated by

$$P_{CMOS} = C_{total} V_0^2 f_{clk} \tag{C.1}$$

where C_{total} is the total capacitance of the circuit, V_0 is the supply voltage, and f_{clk} is the clock frequency of the circuit. It is assumed that the capacitance of multipliers, slicers and lookup tables dominates those of adders and pipelining registers. Second, the critical path (defined as the minimum time required for processing the next new sample [[42], Ch. 3]) is limited by the propagation delay related to the charging and discharging of the CMOS gate and stray capacitances. Thus, the minimum allowed clock period of the processor T_{proc} can be expressed as

$$T_{proc} = \frac{C_{charge} V_0}{k(V_0 - V_t)^2} \tag{C.2}$$

where V_t is the CMOS threshold voltage, and k is a process parameter depending on the material and geometry applied in the CMOS technology [[67], Ch. 2], and C_{charge} is the capacitance to be charged or discharged in a single clock cycle. The interconnect capacitance usually dominates the CMOS gate capacitance, and appears as the major capacitances (e.g. multipliers) within a critical path. For a fair comparison of power consumption between serial processing and parallel processing (with different parallelization level P), we assume that the parallel DD-PR is pipelined with an identical number of registers in each parallel rail as that

of the serial pipelined DD-PR. Please note that fine grain pipelining [[42], p. 69] for multipliers or for lookup tables and bit-level pipelining [[42], p. 482] are not considered here.

For pipelining, the insertion of d pipelining registers reduces the original physical distance within a critical path by d times (where d is the number of pipelining registers as defined in Chapter 3). Thus, the charging capacitance is also reduced by d times because of the reduced coverage of interconnect on the grounded substrate [42], Ch. 6], while the overall capacitance of the circuit is not changed significantly by the addition of pipelining registers. The charging-capacitance reduction allows a faster transition (or a shorter rise-time because of a smaller RC constant), which equivalently reduces the supply voltage from V_0 to βV_0 (for $\beta < 1$) within a clock duration, and the propagation delay $T_{pd,s,pip}$ for pipelined serial processing is

$$T_{pd,s,pip} = \frac{C_{charge,pip}\,(\beta V_0)}{k(\beta V_0 - V_t)^2} \tag{C.3}$$

where $C_{charge,pip}$ is the charging capacitance after pipelining.

For parallel and pipelined processing, the same serial pipelined DD-PR is duplicated in parallel by P times, leading to P-folded increase in the total capacitance in the circuit, while the charging capacitance of each parallel rail, and thus the propagation delay $T_{pd,p,pip}$ remains the same as that of the serial pipelined processing :

$$T_{pd,p,pip} = T_{pd,s,pip} = \frac{C_{charge,pip}\,(\beta V_0)}{k(\beta V_0 - V_t)^2} \tag{C.4}$$

since the processing rate of each parallel rail is P times slower than the original processing rate. Using (C.2) and (C.4)), $T_{pd,p,pip}$ becomes

$$\frac{C_{charge,pip}\,(\beta V_0)}{k(\beta V_0 - V_t)^2} = P\frac{C_{charge}V_0}{k(V_0 - V_t)^2} \tag{C.5}$$

By solving the quadratic equation in β in (C.5) and using (C.1), the power consumption for parallel and pipelined processing can be calculated as

$$P_{CMOS,p,pip} = \frac{C'_{total}\beta^2}{C_{total}P} P_{CMOS} \tag{C.6}$$

Taking $V_0 = 5$ V, $V_t = 0.6$ V, $d = 4$, $C_{charge,pip} = C_{charge}/d$,$C'_{charge} = C_{charge} \times P$, the power consumption for $P = 16$ and $P = 20$ are 2.71% and 2.54% of the original circuit, respectively, leading to a 6.3% reduction in power consumption for performing DD-PR when the parallelization level P increases from 16 to 20.

Annexe D

Reduction of Tracking Bandwidth

This appendix supplements Section 2.5. Intuitively, lower FN-PSD level results in a smaller parallel phase tracking error. In this appendix, we will first discuss the bandwidth reduction of parallel digital feedback loop in more detail based on the implementation form of decision-directed maximum likelihood estimation.

The discrete loop equation for describing parallel DD-MLE is derived based on [75]. Assuming that the phase tracking is performed after perfect equalization and frequency offset compensation, the phasor estimate V [[75], (3)] of parallel DD-MLE in Fig. 2.6 is

$$V = e^{-j\hat{\theta}_{k-i+1}} = \sum_{m=k-(d+1)P+1}^{k-dP} \frac{\hat{d}_m(d_m^* e^{-j\theta_m} + n_m^*)}{\left|\hat{d}_m\right|^2} \tag{D.1}$$

where $\hat{\theta}_{k-i+1}$ is the phase estimate at time index $k-i+1$, i is the parallelization channel index taking from 1 to P, the subscript m refers to time index, θ_m is the true phase, d_m is the complex transmitted QAM signal, n_m is the complex AWGN, $\hat{d}_m$ is the decision symbol of QAM signals.

The LHS of (D.1) corresponds the phase estimates at time indices $k-P+1,\ldots,k$, while the RHS of (D.1) corresponds to the received symbols at time indices $k-(d+1)P+1,\ldots,k-dP$ required to construct the phase estimates. For each round of time interleaving, a common phase estimate is shared by all P channels. For $i = dP,\ldots dP+P-1$, $\hat{\theta}_{k-(d+1)P+1} = \cdots = \hat{\theta}_{k-dP}$. Therefore, for the $(dP)^{th}$ round of time interleaved input signals, the common phase estimate can be factorized out from the summand. Assuming small phase tracking error, such that slicers (decision device) give correct decisions, (D.1) becomes

$$e^{-j\hat{\theta}_{k-i+1}} = e^{-j\hat{\theta}_{k-dP+1}} \sum_{m=k-(d+1)P+1}^{k-dP} (e^{-j(\theta_m-\hat{\theta}_m)} + n_{1,m}^*) \tag{D.2}$$

where the noise term $n_{1,m}^* \triangleq n_m^* e^{j\hat{\theta}_m}/d_m^*$ is rescaled in magnitude by d_m^*, and is rotated by both phase estimate $\hat{\theta}_m$ and the phase of d_m^*. For smaller tracking error, the summand term

in (D.2) becomes

$$\sum_{m=k-(d+1)P+1}^{k-dP}(e^{-j(\theta_m-\hat{\theta}_m)}+n_{1,m}^*)=\sum_{m=k-(d+1)P+1}^{k-dP}[1-j(\theta_m-\hat{\theta}_m)+Re(n_{1,m}^*)+jIm(n_{1,m}^*)] \tag{D.3}$$

For high SNR,the unwrapped phase(D.2) becomes

$$\hat{\theta}_{k-i+1}=\hat{\theta}_{k-dP-i+1}+\frac{1}{P}\sum_{m=k-(d+1)P+1}^{k-dP}(\theta_m-\hat{\theta}_m)-\frac{1}{P}\sum_{m=k-(d+1)P+1}^{k-dP}Im(n_{1,m}^*) \tag{D.4}$$

Note that (D.4) holds for all P channels during each round of time interleaving, as all channels share the same phase estimate. Both the second and the third terms of (D.4) refer to the increment of the discrete loop equation. The second term of (D.4) is the weighted sum of phase errors taken from $k-dP-P+1$ to $k-dP$, with the m^{th} weight equal to the m^{th} received symbol energy. Similarly, the third term of (D.4) is the weighted sum of the imaginary part of $n_{1,m}^*$ taken from $k-dP-P+1$ to $k-dP$ (for high SNR [75]), with the m^{th} weight equal to the m^{th} transmitted symbol energy. These two terms are scaled down by the total symbol energy within the time duration from $dP-P+1$ to $k-dP$.

Annexe E

Frequency Noise Power Spectrum Density

This section supplements Chapters 2 and 3, giving the origin of the f^2-curve at high frequencies in the additive white Gaussian noise (AWGN) FN-PSD in Fig. 2.8. While observable in several references on laser characterization, a thorough analysis is infrequent. Chen et al. [3] attributed such a f^2-curve to optical ASE noise and fiber nonlinearity. However, our previous work in [[40], Fig. 2] experimentally showed the f^2-curve in the absence of fiber transmission contradicting this explanation, and suggesting that the f^2-curve originates from the electrical noise of the coherent receiver. To our best knowledge, Leeson [34] first illustrated that the f^2-curve on FN-PSD is due to white phase noise introduced by thermal noise of electrical oscillators (which appeared as a flat noise floor in PNPSD), and approximated the level of the PN-PSD noise floor using thermal-noise parameters. The small modification for the PN-PSD noise floor can be made by referring to the exact PDF for white phase noise given by Fu and Kam [12].

After coherent detection, the measurement appears as a complex phasor in baseband

$$r_k = Ae^{j\theta_k} + n_k, \tag{E.1}$$

where θ_k is the total laser phase noise, n_k is the zero-mean white Gaussian electrical noise of coherent receiver, with a variance of σ^2 contributed by thermal noise and LO-noise beating, A is the real amplitude of measurement phasor r_k.

The field PSD of r_k consists of a Lorentzian shape near the zero frequency caused by the Brownian phase noise, and a flat floor due to electrical noise covering elsewhere in the field PSD. This two-sided field PSD of white Gaussian noise can be expressed as

$$N_o = \sigma^2 / BW_{2-sided} \tag{E.2}$$

where $BW_{2-sided}$ is the two-sided bandwidth of the spectrum, or the sampling rate of the

ADC (real-time oscilloscope), corresponding to the simulation bandwidth [1]. To obtain PN-PSD, assuming that there is no phase wrapping, we take the angle of E.1

$$Arg\left(r_k\right) = Arg\left[\left(e^{j\theta_k} + n_k/A\right)\right] = \theta_k + \varepsilon_k \tag{E.6}$$

where ε_k is the white phase noise. Applying geometric approach on complex plane for (E.1) shown in [[12], Fig. 1], the exact probability distribution function (pdf) of white phase noise θ depends on measurement, showing a zero-mean Tikhonov distribution with a variance of $\sigma^2/(2|r_k|A)$ [[12], (18)]. For SNR larger than 10 dB, the pdf can be approximated well by a zero-mean Gaussian distribution with a variance of $\sigma^2/(2A^2)$ [[12], (17)], where the factor of $\frac{1}{2}$ can be thought as the equal contribution of in-phase and quadrature-phase noise components. Thus, the phase-noise PSD for white phase noise is given by $N_o/(2A^2)$.

In DSP, we obtain frequency noise by differencing two consecutive phases [[25], (2)], and the FN-PSD, which is the PSD of discretized frequency noise, should have a sinc squared envelope [2] [[25], (4)], but its effect can only be observed at high frequencies around the symbol

1. Ideally, in textbooks such as [48], one usually obtain the PSD level by referring to the magnitude of the autocorrelation of a white noise by using Wiener-Khintchen theorem, such that

$$S(f) = \mathcal{F}\{N_o\delta(\tau)\} = N_o, \tag{E.3}$$

where $\mathcal{F}$ refers to Fourier transform. The ideal noise variance, σ^2_{ideal}, is obtained by setting τ to zero, such that

$$\begin{aligned} R(\tau) &= \int_{-\infty}^{\infty} S(f)e^{j2\pi f\tau}df, \\ \sigma^2_{ideal} &= R(0) = \int_{-\infty}^{\infty} S(f)df. \end{aligned} \tag{E.4}$$

Obviously, σ^2_{ideal} is infinite because of the infinite bandwidth as the upper and lower bounds of the integral. Due to finite measurement bandwidth in practice, i.e., limited by the sampling rate of the ADCs, equation (E.4) becomes

$$\sigma^2 = \int_{-BW_{2-sided}/2}^{BW_{2-sided}/2} N_o df = N_o BW_{2-sided}, \tag{E.5}$$

where σ^2 is the noise variance from measurement or from simulation. It leads to the result in equation (E.2). Note that the PSD level of a white noise, N_o, must be the same in both equation (E.2) and equation equation (E.5). The only difference is that, from our measurement, σ^2 must be finite (and thus is smaller than σ^2_{ideal}) because of the limited bandwidth of instruments. In summary, for theoretical analysis, $N_o = \sigma^2_{ideal}$, while for measurement or for simulation, $N_o = \sigma^2/BW_{2-sided}$, and the values of N_o calculated from both theory and measurement must be the same.

2. As discussed in Appendix B, the sampled (discretized) frequency noise is defined as $f_k = \frac{\theta_{T_s}(kT_s)}{2\pi\ T_s} = \frac{\theta(kT_s+T_s)-\theta(kT_s)}{2\pi\ T_s}$. Using (B.6),

$$f_k = \frac{1}{T_s}\int_{kT_s}^{kT_s+T_s} \Delta f(s)ds, \tag{E.7}$$

where $\Delta f(t)$ is our continuous frequency noise. f_k has been shown to be WSS, and therefore f_k has the same PSD for any k. For simplicity, take $k = 0$,

$$f_o = \int_0^{T_s} h(s)\Delta f(s)ds = h(t) * \Delta f(t), \tag{E.8}$$

rate. Thus, the calculated (discretize) FN-PSD is assumed to be the true (continuous) FN-PSD by ignoring the envelope at frequencies below GHz. The FN-PSD can be simply derived by multiplying f^2 of the phase-noise PSD [3], i.e.

$$S_{\Delta f,AWGN}(f)_{dB} = \left[N_o \Big/ \left(2A^2\right)\right]_{dB} + 20\log_{10} f, \tag{E.10}$$

which agrees with [34]. From (E.10), on the logarithmic FN-PSD, the slope of the f^2-curve does not change as shown in Fig. 3.3 and Fig. 3.4. An increase in electrical noise only brings up the f^2-curve.

where $h(t) = \frac{1}{T_s}$ for $t \in [0, T_s)$ and zero for elsewise, with a Fourier transform $H(f) = e^{-j\pi f T_s} \frac{\sin \pi f T_s}{\pi f T_s}$, and therefore the PSD of f_k, FN-PSD, becomes

$$S_{f_k} = \frac{\sin^2 \pi f T_s}{(\pi f T_s)^2} S_{\Delta f}, \tag{E.9}$$

showing a sinc squared envelope. The 3-dB attenuation of this envelope appears at $f = \frac{1}{4T_s} = \frac{BW_{2-sided}}{4}$, which is well-above 100 MHZ (where the FN-PSD level is used for measuring linewidth of conventional lasers), i.e. $S_{f_k} \approx S_{\Delta f}$ for frequencies below $\frac{BW_{2-sided}}{4}$.

Annexe F

Laser Characterization

Measuring the true laser linewidth requires a coherent receiver (delayed self-homodyne coherent detection) or an interferometric setup with direct detection (p. 180, [7]). In this appendix, we will outline the techniques for linewidth measurement using a coherent receiver, to supplement Chapter 3.

F.1 Measurement Techniques for Laser Linewidth

A coherent receiver captures both real $I(t)$ and imaginary $Q(t)$ components of the optical field, and therefore we completely know the optical field in (B.9) using $A(t) = I(t) + iQ(t)$.

To measure the linewidth using a coherent receiver, we have the following three techniques :

1. The field PSD $S_A(f)$ is found by capturing the complex electrical signal, $A(t)$, computing its autocorrelation, and taking the Fourier transform of the autocorrelation. Then we can fit a Lorentzian curve with a suitable linewidth to the measurement.

   This method is only suitable for large laser linewidths (a few MHz or higher) because small-linewidth (below MHz) lasers give sharper Lorentzian field PSD which requires a higher FFT resolution to identity the 3-dB bandwidth of the field PSD.

   Higher FFT resolution requires a longer observation window in time. However, usually low-linewidth laser sources contain unexpected control tones. Together with the low-frequency noise such as flicker noise (due to mechanical vibration), the Lorentzian field PSD peak moves around zero frequency, resulting in a much larger estimated linewidth.

2. Another method is to compute the variance of phase noise increment captured with the coherent setup and use (B.8). The linewidth can be estimated using [(6), [55]] :

$$\Delta\upsilon = 2\pi\sigma^2_{\Delta f} = 2\pi\left[\frac{\sigma^2_{\theta_\tau}}{(2\pi)^2\tau}\right] = \frac{\sigma^2_{\theta_\tau}}{2\pi\tau} = \frac{E\left[|\theta(t+\tau)-\theta(t)|^2\right]}{2\pi\tau}. \tag{F.1}$$

3. We can also generate a frequency noise (FN) -PSD, where frequency noise can be taken as the consecutive phase difference divided by $2\pi\tau$. The flat region of the PSD equals the variance of frequency noise $\sigma^2_{\Delta f}$.

The disadvantage of the last two methods (2), (3) is that the white phase noise [3] from the electrical components or the ASE noise from EDFA will lead to over-estimating the laser linewidth when considering the frequency region beyond 100 MHz (Fig. 3 in [3]).

In the absence of white phase noise, the last two methods can be applied, provided that the phase noise is a Wiener process which gives a flat frequency noise PSD.

F.2 Experimental Setup

To measure the linewidth of TeraXion laser, we performed an experiment using a delayed self-homodyne coherent detection setup shown in Fig. F.1. In the setup, a 3 dB coupler was

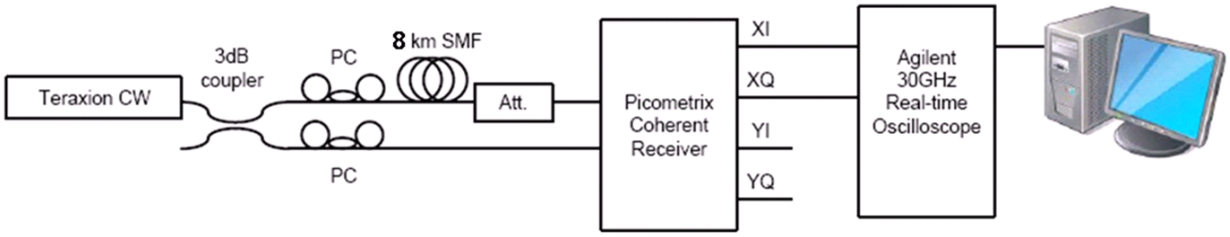

FIGURE F.1 – Experimental setup of the delayed self-homodyne coherent detection.

used to split the CW laser equally into a local oscillator (LO) and a signal (Sig). Note that both LO and Sig were from the same source, and we refer to them differently for convenience only. The polarization states of both optical fields were controlled by polarization controllers. The signal and the LO were decorrelated by an 8 km-long standard single-mode fiber that introduced a sufficient time delay on the signal [1] To approach the real coherent system where Sig is always lower than LO, we attenuated the Sig. The Sig was further lowered using a 5 dB optical attenuator to make sure that the optical power of the LO, $P_{LO}(t)$, is much higher than that of the Sig, $P_{Sig}(t)$, i.e. $P_{LO}(t) \gg P_{Sig}(t)$. Fiber attenuation, connector insertion loss and the overall insertion loss of the optical attenuator introduces approximately 10 dB on Sig. The photocurrents of I and Q in the x polarization are (p. 68, [21]) :

$$i_I(t) \sim R_p \sqrt{P_{LO}(t) P_{Sig}(t)} \cos\left[2\pi (f_{LO} - f_{Sig}) t + \theta_{CW}(t + \tau_o) - \theta_{CW}(t)\right] \tag{F.2}$$

$$i_Q(t) \sim R_p \sqrt{P_{LO}(t) P_{Sig}(t)} \sin\left[2\pi (f_{LO} - f_{Sig}) t + \theta_{CW}(t + \tau_o) - \theta_{CW}(t)\right], \tag{F.3}$$

1. The time delay between LO and Sig should be longer than the coherence time of the CW source, explained by equation (5.22) in p. 186 of [7].

where $\theta_{CW}(t)$ is the phase noise of the laser under test, f_{LO} and f_{Sig} are the center frequency of LO and Sig, respectively, R_p is the responsivity of each photodetector (assuming that the two photodetectors are identical). Since LO and Sig were decorrelated by a sufficient time, τ_o, $\theta_{CW}(t)$ and $\theta_{CW}(t+\tau_o)$ can be considered as two independent random phase noises, and therefore the measured phase noise increment (refer to (B.8)) has a variance twice of the original variance.

During the experiment, the CW laser had a constant output power of 15 dBm. Because of the insertion losses due to the 3 dB couplers, polarization controllers, single mode fiber (SMF), optical attenuator as well as connector losses, the $P_{LO}(t)$ and $P_{Sig}(t)$ arriving the coherent receiver were 9.5 dBm and 0 dBm, respectively.

We used a 25 GHz-bandwidth Picometrix coherent receiver to detect the inphase (I) and quadrature-phase (Q) components of the CW field, and convert them into the complex RF signals. An Agilent DSA 930004L real-time oscilloscope was used to undergo the analog-to-digital conversion (ADC) at a sampling rate of 20 Gsa/s. We adjusted two polarization controllers to maximize the RF signals shown on the scope.

Finally, the data was sent to a computer for offline digital signal processing (DSP). For delayed self-homodyne coherent detection, f_{LO} and f_{Sig} can be assumed to be the same, as long as the laser is stably locked to its center frequency. This is the case for the TeraXion CW laser, as we verified experimentally.

To measure the laser linewidth, the FN-PSD shown in Fig. 3.4 in Chapter 3 was generated using method 3 previously mentioned. The following DSP was used :

1. Form a complex number vector based on the data from I and Q ;
2. Take the phase of the complex vector using the Matlab function *angle* ;
3. Generate the phase noise increment by taking the difference between successive phases using the Matlab function *diff*, and divide the increment by 2 *pi* ;
4. Fourier-transform the phase-noise-increment vector. The FFT size is 2×10^6, equal to the length of the vector, which corresponds to a frequency resolution of $\frac{\text{Sampling rate}}{2\times10^6} =$ 10 kHz, and then take the absolute square of the vector, multiplied with scaling factors to form a PSD, as mentioned inside the definition of PSD in Matlab ;
5. Take an average of the FN-PSD over 100 new measurements. There should not be any moving average across the frequency axis (spectrum) to avoid distortion due to windowing.

Then we can calculate the laser linewidth based on the FN-PSD level between 10 and 100 MHz.

Annexe G

Clock Tone and Timing Phase Estimation

This appendix provides background on timing phase estimation and its failure due to the polarization effects of an optical channel. We introduce the problem with a brief literature review. We will look into a well-known frequency-domain timing phase estimation. Finally, we will see the impact of optical channel on clock tone magnitude. This appendix is to supplement Chapter 6 and Chapter 7.

G.1 Introduction

Before the start of digital signal processing (DSP), the clock-tone (CT) magnitude is used for chromatic dispersion (CD) estimation [57]. Timing phase estimation makes use of the CT phase to determine the timing phase during resampling [19]. Timing phase estimation fails for certain SOP (state of polarization) because the CT magnitude is largely suppressed at half-symbol delay, or half-baud differential group delay (DGD) [19], [20]. Loss of CT can be avoided by pre-rotating the received SOP to avoid equal power splitting of dual-polarization signals between X and Y-polarizations [59, 60, 61, 62]. Different algorithms were proposed to rotate the SOP to avoid CT loss, but they involve sophisticated algorithms at the symbol rate or higher using substantial numbers of block-wise multipliers [59, 60, 61, 62]. SOP rotation is currently mostly confined to CT extraction, and is power hungry and increases implementation complexity. Please note that CT magnitude does not infer directly to CT phase. Even if CT magnitude is strong enough, the CT phase may still be distorted due to both DGD and SOP rotation.

G.2 Timing Phase Estimation (TPE)

Recently FD-Godard TPE gains its popularity in optical coherent communications : First, DSP in frequency domain is proved to save more hardware complexity. Second, chromatic dispersion compensation (CDC) is first performed at the beginning of the DSP in frequency domain (FD) [1]. TPE can then be directly applied after FD-CDC instead of after MIMO (polarization demultiplexing), and feed the timing information back to the ADCs for adjusting the sampling instants. This can reduce the feedback delay and thus the timing jitter.

The Godard timing phase estimation employs the phase of the clock tone. The clock tone is defined as the complex autocorrelation function (ACF) of the received signal **in frequency domain**, $r_i(f)$, of either X- or Y-polarization ($i = x, y$) at clock frequency. The complex ACF for X-polarization is defined as

$$C_{ii}(F) = \int r_i(f) r_i^*(f+F) df \tag{G.1}$$

where F is the shift parameter of the ACF. The clock tone refers the complex value of the ACF at the clock frequency F_{CT} :

$$F_{CT} = \pm N_{FFT}(1 - \frac{R_s}{R_{ADC}}), \tag{G.2}$$

where R_{ADC} is the sampling rate of ADC, while R_s is the symbol rate of the received signal. Therefore, the CT magnitude and CT phase (in radian) are $|C_{ii}(F_{CT})|$ and $\angle\Big(C_{ii}(F_{CT})\Big)$, respectively. The timing offset normalized to symbol period is given by $\angle\Big(C_{ii}(F_{CT})\Big)/2\pi \in [-0.5, 0.5)$, in unit interval (U.I.).. Note that timing phase and timing offset usually refer to the same meaning in our context. It is obvious that the magnitude of the clock tone, $|C_{ii}(F)|$, must be well-above the noise floor, or the corresponding phase will be noisy that results in timine phase jitter.

G.3 Impact of Polarization Effects on TPE

Hauske et al. [19] first analysed the polarization effects of optical channels on signal clock tones. This subsection is to revisit Hauske et al.'s work to provide a background to understand timing phase recovery mentioned in Chapter 5, 6 and 7. Assuming that the CD is compensated before the timing phase recovery, the polarization effects, SOP rotation and DGD, can be described using the Jones matrix in frequency domain in (5.1) in Chapter 5 (also appeared as equation (2) in [19]). Let $\underline{s}(f) = [s_x(f), s_y(f)]^T$ be the transmitted dual-polarization (DP) signal in frequency domain, the receive signal $\underline{r}(f) = [r_x(f), r_y(f)]^T$ becomes

$$\underline{r}(f) = H(f)\underline{s}(f). \tag{G.3}$$

1. Unlike the time-domain CDC in which the FIR filters require 2 samples per symbol, frequency-domain CDC can be applied less than 2 samples per symbol, effectively reducing the hardware complexity of coherent receivers.

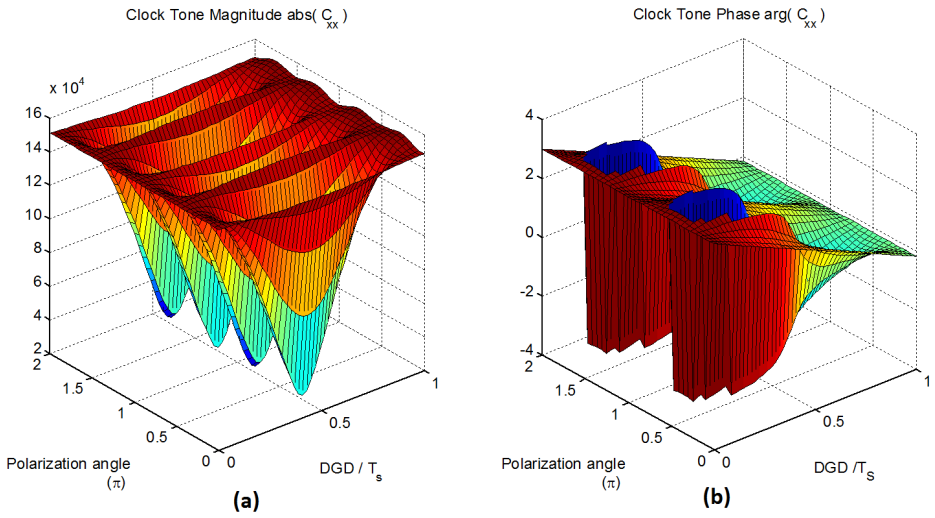

FIGURE G.1 – (a) Clock tone magnitude and (b) clock tone phase of received signal either on X-polarization under the effect of polarization angle and DGD. The receiver bandwidth is 0.7 Baud. A.u. : arbitrary unit.

The authors in [19] started with the assumption that the sampling offsets of X-ADCs and Y-ADCs are the same. In our simulation, we used zero sampling offset. The results of CT magnitude and CT phase based on the autocorrelation function of the received signal (in frequency domain) in X-polarization, C_{xx}, are shown in Fig. G.1a and Fig. G.1b, respectively. Note that the results using C_{yy} are not shown here because it shows similar results as mentioned in [19]. In Fig. G.1a, the clock tone magnitude drops significantly at $\theta = \pi/4$ and θ is equal to half symbol rate T_s, where θ is the polarization rotation angle and τ is the time delay between two polarization channels caused by DGD. Noise buries this low clock magnitude, and therefore possibly affects the calculated phase. In Fig. G.1b, the CT phase, which is used to estimate timing phase, is a function of θ and τ. This indicates that the timing phase estimated from the Godard TPE is affected by the polarization effects of optical channel. Hauske et al. [19] showed that ϕ does not affect the clock tone.

Annexe H

Stokes Space Polarization Demultiplexing

This appendix provides a literature review of polarization demultiplexing using Stokes space parameters, and supplements our argument in Chapter 6 that MIMO-FIRs are necessary for suppressing polarization crosstalk because of the inadequate performance using Stokes space polarization demultiplexing.

Principle

Conventional polarization demultiplexing (PolDemux), uses sample-wise MIMO-FIRs adapted via the constant modulus algorithms (CMA) or a least-mean squares (LMS) approach [8].

Szafraniec et al. [63, 64] proposed a polarization demultiplexing based on Stokes parameters (SS-PolDemux). The signal representation of M-ary QAM (quadrature amplitude modulation) in Stokes space, as a function of three Stokes parameters (s_1, s_2, s_3), can be traced as a lens-shaped object (shown in Fig. 1 and Fig. 2 in [64], and our Fig. 6.2) on the Poincaré sphere, inspiring SS-PolDemux. Szafraniec, et al., applied a plane fitting to obtain the 3-dimensional (D) Stokes parameters of the normal of this lens-shaped object to estimate the polarization orientation angle (azimuth) and the phase delay angle (ellipticity) between two polarizations using Stokes parameters. The estimated values of these two angles are then used to calculate a 2×2 Jones-space transfer matrix that represents the inverse of the optical channel as shown in (5.1). The polarization-coupled signal is derotated by this Jones matrix to decouple the two polarization channels, i.e., a sample-wise multiplication between the received samples and the Jones matrix.

As SS-PolDemux is not affected by sampling phase, carrier frequency offset or laser phase noise, this technique can be applied directly before the DSP stages such as retiming and MIMO.

On the other hand, to date the SS-PolDemux has been limited to metrology applications because of its heavy computational effort in calculating Stokes parameters via a plane fitting, which requires sufficiently long averaging. Hence, there is a tradeoff between the accuracy of estimation and the speed of polarization rotation. Besides, the lens-shaped object becomes more like a sphere for half-symbol misalignment between X and Y polarization at the transmitter for zero DGD, half-symbol differential group delay at 45^o (orientation angle) polarization rotation and limited receiver bandwidth, especially for higher-order QAM signals [1].

Literature Review

Schmogorow et al. [51, 11] proposed a simplified gradient search algorithm based on s_0 alone (the total power of DP signals) for SS-PolDemux to reduce the computational complexity for real-time processing. This technique was verified only by a back-to-back experiment without DGD-impaired transmission, together with Agilent OMA software that provides additional compensation of the residual polarization crosstalk [64]. It is well known that, for long haul transmission [36], large PMD requires more MIMO-FIR cross taps. A pure SS-PolDemux cannot compensate the DGD effect because of the absence of FIR filters.

On the other hand, Muga and Pinto [38] recently proposed a computationally-costly sample-by-sample adaptive 3D SS-PDM algorithm to track the orientation of the normal component of the signal representation in Stokes space. Again, however, their pure simulation work ignored the DGD introduced during transmission [2]. Other than the DGD effect, the simulation work of Muga and Pinto did not take into account the imperfect sampling instants introduced by receiver ADCs (analog-to-digital converters), which definitely requires FIR taps for timing phase recovery.

MIMO is necessary

For short reach systems, small DGD can induce significant OSNR penalty for high symbol rate signals in the absence of DGD compensation [28]. Even with zero DGD, Winzer et al. [28, 70, 68] indicated that polarization crosstalk as low as -25 dB [3] still requires MIMO processing to achieve OSNR power penalty below 0.5 dB. In the presence of noise and distortion caused by limited receiver electrical bandwidth and polarization leakage of non-ideal polarization beam splitters and optical hybrid, SOP estimates in Stokes space are not accurate enough to totally suppress the polarization crosstalk.

1. In our experiment and simulation, both a 32 Gbaud DP-QPSK and 16 Gbaud DP-16QAM require at least 16GHz to guarantee a good lens object on the Poincaré sphere.
2. Muga and Pinto also put a decent value of receiver bandwidth in their simulation work. However, for limited receiver bandwidth (e.g. 32 Gbaud signal with only 16 GHz receiver bandwidth), the polarization tracking using Stokes-space algorithms can never be accurate.
3. The polarization crosstalk is defined as the ratio of crosstalk power-to-signal power in dB.

Conclusion

The important conclusion is that a pure SS-PolDemux (without MIMO) with moderate SOP estimation accuracy cannot completely compensate polarization crosstalk, and a pure SS-PolDemux with very high SOP estimation accuracy cannot provide any DGD compensation. In short, MIMO-FIRs are indispensable for reducing the penalty caused by polarization crosstalk. Yu et al. [74, 73], recognized that Stokes space parameters can be used to initialize MIMO-FIR taps to avoid long MIMO convergence and singularity problem, at the expense of extra circuitry for the SOP estimation. DGD compensation was conserved as MIMO-FIRs were still used.

Annexe I

Experimental Generation of Random SOPs

In this appendix, we present the experimental technique for generating random SOPs uniformly distributed over the Poincaré sphere. Each SOP is represented by a 3-dimensional (D) Stokes vector, i.e., s_1, s_2 and s_3 normalized to give a unit radius. One may suggest the use of spherical coordinates and generate random angles. However, as spherical coordinates are parameterized by two angles and a radius, the resultant random 3-D vectors can never cover the sphere uniformly, i.e. the two poles of the Poincaré sphere are crowded with random points. Therefore, a random (uniformly distributed) spherical sampling is necessary.

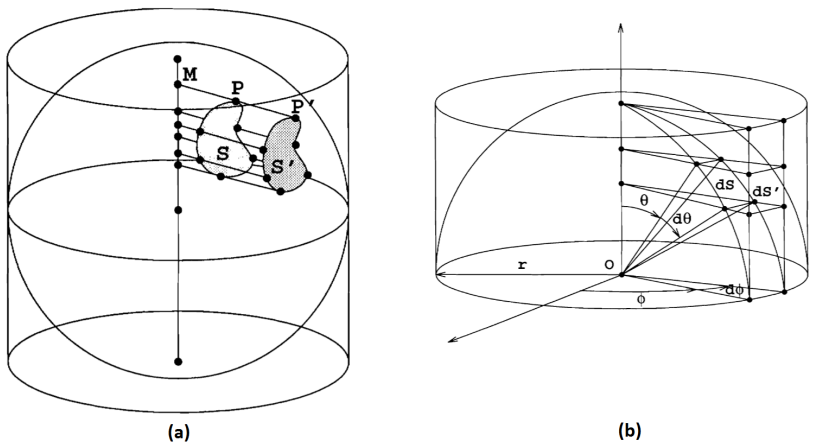

FIGURE I.1 – (a) A sphere and a right cylinder circumscribed about the sphere. Spherical points P and area S and the corresponding axial projections P' and S' on the cylindrical surface. (b) Differential area dS on the sphere and its axial projection dS' on the cylinder.

In view of this, we employed the method using the principle of Archimedes theorem (local)

suggested by Shao and Badler [54] in 1996, where axial projection of any measurable region on a sphere onto the right circular cylinder circumscribed about the sphere preserves area.

Refer to Fig. I.1a. Suppose P is a point on the sphere and M, a point on the axis of the circumscribed cylinder, is the closest point to P. Extrapolate the line MP to intersect the surface of the cylinder at point P'. The point P' is called the axial projection of the point P on a spherical surface onto the surface of the right circular cylinder circumscribed about the sphere. The above theorem means that the area S' is equal to the area S. Mathematically, one has to prove that the two differential areas, dS and dS' are equal to each other shown in Fig. I.1b. For details, please refer to [54].

As random sampling (with uniform distribution) can be easily done on a cylindrical surface, we can generate a random point on the cylinder [-1, 1]×[0, 2π], and then find its inverse axial projection on the Poincaré sphere. For random sampling on a surface, please refer to Leon-Garcia's textbook [14]. As a result, the random spherical sampling avoids the crowded random points on the two poles of the Poincaré sphere.

In experiment, we could only control five waveplates (4 quarter waveplates with a half waveplates at the middle) but not the 3-D Stokes parameters inside the polarization synthesizer (Agilent N7786B), and a calibration (mapping between waveplates' positions and 3-D Stokes parameters) was necessary. Since the N7786B is not able to measure the Stokes parameters of DP-QPSK signals (which behaves as if a unpolarized light due to random data), only a continuous wave without modulation was used for calibration. We adjusted only 3 waveplates in N7786B to cover the Poincaré sphere as densely as possible, and recorded the corresponding Stokes parameters in a look-up table. Thus, the randomly generated stokes parameters can be used to sweep SOPs using the N7786 via our own look-up table.

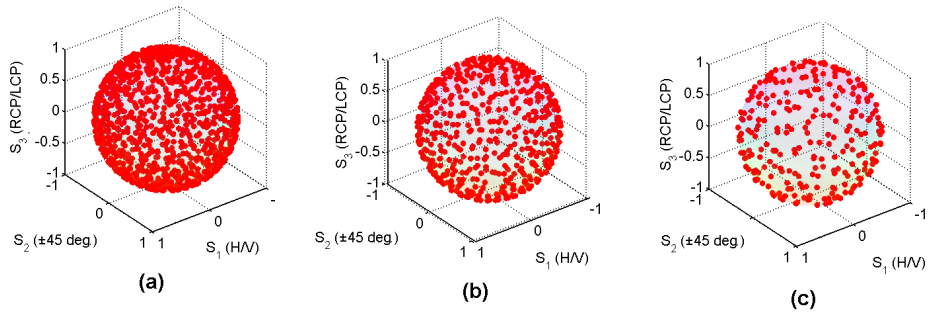

FIGURE I.2 – The results of random spherical sampling showing (a) 1000 random SOPs ; (b) 500 random SOPs (interleaved from (a)) ; (c) 250 random SOPs (interleaved from (a)).

Annexe J

Publication List

The followings are my publications since 2010.

1. **W. C. Ng**, A. T. Nguyen, C. S. Park, and L. A. Rusch, "Enhancing Clock Tone via Polarization Pre-rotation : A Low-complexity, Extended Kalman Filter-based Approach," submitted to *Optical Fiber Communication Conference*, March, 2015.
2. **W. C. Ng**, A. T. Nguyen, C. S. Park, and L. A. Rusch, "MIMO-FIR Tap Reduction via Polarization Pre-rotation for 100 Gbps Short-Reach Applications," submitted to *IEEE J. Lightw. Technol.*
3. **W. C. Ng**, A. T. Nguyen, C. S. Park, and L. A. Rusch, "Reduction of MIMO-FIR Taps via SOP-Estimation in Stokes Space for 100 Gbps Short Reach Applications," *European Conference and Exhibition on Optical Communication*, September, 2014, P3.3.
4. A. T. Nguyen, **W. C. Ng**, C. S. Park, and L. A. Rusch, "An optimized 16-QAM constellation for mitigating impairments of phase noise and limited transmitter ENOB in optical coherent detection systems," *European Conference and Exhibition on Optical Communication*, September, 2014, Tu.1.3.5.
5. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch, "Overcoming Phase Sensitivity in Real-time Parallel DSP for Optical Coherent Communications : Optically Filtered Lasers," *IEEE J. Lightw. Technol.*, **Vol**. 32, **No**. 3, pp. 411 - 420, Feb. 2014. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6678710
6. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch, "Impact of Sinusoidal Tones on Parallel Decision-Directed Phase Recovery for 64-QAM," *IEEE Photonics Letters Technology*, **Vol**. 26, **No**. 5, pp. 486 - 489, Mar. 2014. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=6701345
7. T. N. Huynh, A. T. Nguyen, **W. C. Ng**, L. Nguyen, L. A. Rusch, and L. P. Barry, "BER Performance of Coherent Optical Communications Systems Employing Monolithic Tunable Lasers With Excess Phase Noise," *IEEE J. Lightw. Technol.*, **Vol**. 32, **No**. 10, pp. 1973 - 1980 May. 2014.

8. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch, "Parallel and Pipelined Decision-Directed Phase Recovery for 64-QAM in the Presence of Sinusoidal Tones," *Optical Fiber Communication Conference*, M2A.4, March, 2014.
9. B. Filion , **W. C. Ng**, An T. Nguyen, L. A. Rusch and S. LaRochelle, "Wideband wavelength conversion of 16 Gbaud 16-QAM and 5 Gbaud 64-QAM signals in a semiconductor optical amplifier, " *Optics Express*, Vol. 21, Issue 17, pp. 19825-19833 ,2013.
10. B. Filion , **W. C. Ng**, An T. Nguyen, L. A. Rusch and S. LaRochelle, "Wideband Wavelength Conversion of 5 Gbaud 64-QAM Signals in a Semiconductor Optical Amplifier," *European Conference and Exhibition on Optical Communication*, P.3.19, September, 2013.
11. **W. C. Ng**, T. N. Huynh, A. T. Nguyen, S. Ayotte, C. S. Park, L. P. Barry and L. A. Rusch, "Improvement due to Optically Filtered Lasers in Parallel Decision-Directed Phase Recovery for 16-QAM," *The 10th Conference on Lasers and Electro-Optics Pacific Rim, and The 18th OptoElectronics and communications Conference / Photonics in Switching 2013*, WR3-2, June 2013.
12. T. Huynh, **W. C. Ng**, A. T. Nguyen, L. Nguyen, L. Rusch, and L.Barry, "Tracking Excess Noise from a Monolithic Tunable Laser in Coherent Communication Systems," *Optical Fiber Communication Conference 2013*, JW2A.46, March, 2013.
13. **W. C. Ng**, A. T. Nguyen, S. Ayotte, C. S. Park, L. A. Rusch, "Optically-filtered Laser Source Enabling Improved Phase Tracking in Coherent Transmission Systems," *European Conference and Exhibition on Optical Communication*, P.3.15, September, 2012.
14. K. Dolgaleva, **W. C. Ng**, , L. Qian, J. S. Aitchison, M. Carla Camasta and Mark Sorel, "Compact Highly-Nonlinear AlGaAs Waveguides for Efficient Wavelength Conversion, " *Optics Express*, Vol. 19, pp. 12440-12455, 2011.
15. **W. C. Ng**, Q. Xu, K. Dolgaleva, S. Doucet, D. Lemus, P. Chretien, W. Zhu, L. A. Rusch, S. LaRochelle, L. Qian, J. S. Aitchison, "Error-free 0.16pi-XPM-based All-Optical Wavelength Conversion in a 1-cm-long AlGaAs waveguide, " *Proc. IEEE Photonics Society Annual Meeting 2010*, pp. 54-55 , November, 2010.
16. K. Dolgaleva, **W. C. Ng**, L. Qian, and J. S. Aitchison, "Efficient Self-Phase Modulation, Cross-Phase Modulation, and Four-Wave Mixing in AlGaAs Waveguides, " *Frontier in Optics 2010*, FTuC5, October, 2010.
17. K. Dolgaleva, **W. C. Ng**, L. Qian, J. S. Aitchison ,M. C. Camasta and M. Sorel, "Broadband Self-Phase Modulation, Cross-Phase Modulation, and Four-Wave Mixing in 9-mm-long AlGaAs Waveguides, " *Optics Letters*, Vol. 35, Issue 24, pp. 4093-4095, 2010.
18. A. D. Simard, N. Ayotte, Y. Painchaud, S. Bédard, and S. LaRochelle, "Impact of Sidewall Roughness on Integrated Bragg Gratings, " *IEEE J. Lightw. Technol.*, **Vol**. 29, **No**. 24, pp. 3693 - 3704, Dec. 2011. (Personal tutoring [1] to A. D. Simard.)

1. Contributions on methodology and on derivation for equations (15) – (26), Appendix A and B

Bibliographie

[1] Govind P. Agrawal. *Nonlinear Fiber Optics*. Academic Press, Inc., 2001.

[2] S. Ayotte, F. Costin, G. Brochu, M. J. Picard, A. Babin, X. Liu F. Pelletier, and S.Chandrasekhar. White Noise Filtered C-band Tunable Laserfor Coherent Transmission Systems. In *Optical Fiber Communication Conference*, page OTu1G5, 2012.

[3] X. Chen, A. A. Amin, and W. Shieh. Characterization and Monitoring of Laser Linewidths in Coherent Systems. *IEEE Journal of Lightwave Technology*, 29(17) :2533–2537, September 2011.

[4] D. Crivelli, M. Hueda, H. Carrer, and J. Zacha. A 40nm CMOS single-chip 50Gbps DPQPSK or BPSK transceiver with electronic dispersion compensation for coherent optical channels. In *International Solid-State Circuits Conference*, pages 328–330, Feburary 2012.

[5] D. E. Crivelli, M. R. Hueda, H. S. Carrer, M. del Barcoand R. R. Lopez, P. Gianni, J. Finochietto, N. Swenson, P. Voois, and O. E. Agazzi. Architecture of a Single Chip 50 Gbps DP QPSK BPSK Transceiver with Electronic Dispersion Compensation for Coherent Optical Channels. *IEEE Transactions on Circuits and Systems 1*, 61(4) :1012, March 2014.

[6] J.-M. Delgado-Mendinueta, B. J. Puttnam, R. S. Luis, S. Shinada, and N. Wada. Fast Equalizer Kernel Initialization for Coherent PDM-QPSK Burst-mode Receivers Based on Stokes Estimator. In *Signal Processing in Photonic Communication*, page SPM2E.2, Rio Grande, Puerto Rico United States, July 2013.

[7] D. Derickson. *Fiber optics test and measurement*. Prentice Hall PTR, 1998.

[8] I. Fatadin, D. Ives, and S. J. Savor. Laser linewidth tolerance for 16-QAM coherent optical systems using QPSK partitioning. *IEEE Photonics Technology Letters*, 22 :631–633, May 2010.

[9] I. Fatadin, D. Ivex, and S. J. Savory. Blind Equalization and Carrier Phase Recovery in a 16-QAM Optical Coherent System. *IEEE Journal of Lightwave Technology*, 27(15) :3042–2049, August 2009.

[10] B. Filion, W. C. Ng, A. T. Nguyen, L. A. Rusch, and S. LaRochelle. Wideband Wavelength Conversion of 5 Gbaud 64 QAM Signals in a Semiconductor Optical Amplifier. In *European Conference and Exhibition on Optical Communication*, page P.3.19, London, September 2013.

[11] Wolfgang Freude, Rene Schmogrow, Philipp C. Schindler, Stefan Wolf, Bernd Nebendahl, Christian Koos, and Juerg Leuthold. Polarisation Demultiplexing in Coherent Receivers with Real-Time Digital Signal Processing. In *Transparent Optical Networks (ICTON)*, 2013.

[12] H. Fu and P. Y. Kam. Exact phase noise model and its application to linear minimum variance estimation of frequency and phase of a noisy sinusoid. In *IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2008.

[13] Y. Gao, A. P. T. Lau, C. Lu, J. Wu, Y. Li, X. Kun, W. Li, and T. J. Lin. Low-Complexity Two Stage Carrier Phase Estimation for 16 QAM Systems using QPSK Partitioning and Maximum Likelihood Detection. In *Optical Fiber Communication Conference*, page OMJ6, 2011.

[14] Alberto Leon Garcia. *Probability and Random Processes for Electrical Engineering*. Pearson Education, 2008.

[15] R. D. Gaudenzi, T. Garde, and V. Vanghi. Performance anaylsis of decision directed maximum likelihood phase estimators for M PSK modulated signals. *IEEE Transactions on Communication*, 43(12) :3090–3100, December 1995.

[16] P. Gianni, G. Corral Briones, H. Carrer, and M. Hueda. Compensation of laser frequency fluctuations and phase noise in 16-QAM coherent receivers. *IEEE Photonics Technology Letters*, 25(5) :442–445, March 2013.

[17] P. Gianni, G. Corral Briones, C. Rodriguez, H. Carrer, and M. Hueda. A new parallel carrier recovery architecture for intradyne coherent optical receivers in the presence of laser frequency fluctuations. In *IEEE Global Telecommunications Conference*, pages 1–6, December 2011.

[18] A. H. Gnauck, P. Winzer, A. Konczykowska, F. Jorge, J. Y. Dupuy, M. Riet, G. Charlet, B. Zhu, and D. W. Peckham. Generation and Transmission of 21.4 Gbaud PDM 64 QAM using a High Power DAC Driving a Single I Q modulator. *IEEE Journal of Lightwave Technology*, 30(4) :532–536, 2012.

[19] F. N. Hauske, N. Stojanovic, C. Xie, and M. Chen. Impact of Optical Channel Distortion to Digital Timing Recovery in Digital Coherent Transmission Systems. In *International Conference on Transparent Optical Networks*, page We.D1.4, 2010.

[20] F. N. Hauske, C. Xie, N. Stojanovic, and M. Chen. Analysis of Polarization Effects to Digital Timing Recovery in Coherent Receivers of Optical Communication Systems. In *Photonische Netze*, page P6, 2010.

[21] Keang Po Ho. *Phase modulated optical communication systems*. Springer, 2005.

[22] L. Huang, D. Wang, A. P. T. Lau, C. Lu, , and S. He. Performance analysis of blind timing phase estimators for digital coherent receivers. *Optics Express*, 22(6) :6749–6763, March 2014.

[23] E. Ip and J. Kahn. Feedforward Carrier Recovery for Coherent Optical Communications. *Journal of Lightwave Technology*, 25(9) :2675–2692, September 2007.

[24] Steven M. Kay. *Fundamentals of Statistical Signal Processing Estimation Theory*. Prentice Gall Signal Processing Series. Pearson, 1993.

[25] K. Kikuchi. Characterization of semiconductor-laser phase noise and estimation of bit-error rate performance with low-speed offline digital coherent receivers. *Optics Express*, 20(5) :5291–2302, Feb 2011.

[26] K. Kikuchi. Clock Recovering Characteristics of Adaptive Finite Impulse Response filters in Digital Coherent Optical Receivers. *Optics Express*, 19(6) :5611–5619, March 2011.

[27] K. Kikuchi. Performance Anaylsis of Polarization Demultiplexing based on Constant-Modulus Algorithm in Digital Coherent Receiver. *Optics Express*, 19(10) :9868, 2011.

[28] H. Kogelnik and P. Winzer. PMD Outage Probabilities Revisited. In *Optical Fiber Communication Conference*, page OTuN3, 2007.

[29] M. Kuschnerov, T. Bex, and P. Kainzmaier. Energy efficient Digital Signal Processing. In *Optical Fiber Communications Conference*, page Th3E.7, San Francisco, CA, USA, 2014.

[30] M. Kuschnerov, K. Piyawanno, M. Alfiad, B. Spinnler, A. Napoli, and B. Lankl. Impact of mechanical vibrations on laser stability and carrier phase estimation in coherent receivers. *IEEE Photonics Technology Letters*, 22(15) :1114–1116, August 2010.

[31] M. Kuschnerov, K. Piyawanno, M.S. Alfiad, B. Spinnler, A. Napoli, and B. Lankl. DGD-tolerant Timing Recovery for Coherent Receivers. In *OptoElectronics and Communications Conference*, pages 7B1–4, Sopporo Convention Center, Japan, July 2010.

[32] A. P. T. Lau, T. S. R. Shen, W. Shieh, and K.-P. Ho. Equalization enhanced phase noise for 100Gbps transmission and beyond with coherent detection. *Optics Express*, 18(16) :17239–17251, 2010.

[33] D. Lavery, R. Maher, D. Millar, B. C. Thomsen, P. Bayvel, and S. Savory. Demonstration of 10 Gbit/s colorless coherent PON incorporating tunable DS-DBR lasers and low-complexity parallel DSP. In *Optical Fiber Communication Conference*, page PDP5B.10, Los Angeles, California United States, March 2012.

[34] D. B. Leeson. A simple model of feedback oscillator noise spectrum. In *IEEE*, volume 54, pages 329–330, February 1966.

[35] M. Li, N. Deng, F. N. Hauske, Q. Xue, X. Shi, Z. Feng, S. Cao, and Q. Xiong. Optical Burst-mode Coherent Receiver with a Fast Tunable LO for Receiving Multi-wavelength Burst Signals. In *Optical Fiber Communication Conference*, page OTu1G, 2012.

[36] N. C. Mantzoukis, C. S. Petrou, A. Vgenis, T. Kamalakis, I. Roudas, and L. Raptis. Outage Probability due to PMD in Coherent PDM QPSK Systems with Electronic Equalizer. *IEEE Photonics Technology Letters*, 22(16) :1247–1249, August 2010.

[37] C. J. McKinstrie, H. Kogelnik, R. M. Jopson, S. Radic, and A. V. Kanaev. Four wave mixing in fibers with random birefringence. *Optics Express*, 12(10) :2033–2055, 2004.

[38] N. J. Muga and A. N. Pinto. Adaptive 3D Stokes Space Based Polarization Demultiplexing Algorithm. *IEEE Journal of Lightwave Technology*, 32(19) :3290–3298, 2014.

[39] W. C. Ng, T. N. Huynh, A. T. Nguyen, S. Ayotte, C. S. Park, L. P. Barry, and L. A. Rusch. Improvement due to Optically Filtered Lasers in Parallel Decision-Directed Phase Recovery for 16-QAM. In *The 10th Conference on Lasers and Electro-Optics Pacific Rim, and The 18th OptoElectronics and communications Conference / Photonics in Switching 2013*, page P3.15, June 2013.

[40] W. C. Ng, A. T. Nguyen, S. Ayotte, C. S. Park, and L. A. Rusch. Optically-filtered Laser Source Enabling Improved Phase Tracking in Coherent Transmission Systems. In *European Conference and Exhibition on Optical Communication*, page P3.15, September 2012.

[41] M. Oerder and H. Meyr. Digital Filter and Square Timing Recovery. *IEEE Transactions on Communications*, 36(5) :605 – 612, May 1988.

[42] Keshab K. Parhi. *VLSI Digital Signal Processing Systems : Design and Implementation.* ISBN : 978-0-471-24186-7. Wiley, December 1998.

[43] W.-R. Peng, T. Tsuritani, and I. Morita. Transmission of highbaud PDM 64QAM signals. *IEEE Journal of Lightwave Technology*, 31(13) :2146–2162, July 2013.

[44] T. Pfau, S. Hoffmann, and R. N. Hardware-Efficient Coherent Digital Receiver Concept With Feedforward Carrier Recovery for M-QAM Constellations. *IEEE Journal of Lightwave Technology*, 27(8) :989–999, 2009 2009.

[45] F. Pittala, A. Mezghani, I. Slim, and J. A. Nossek. Low-complexity training-aided 2 by 2 MIMO frequency domain fractionally-spaced equalization. *Optical Fiber Communications Conference*, page Th2A.42, March 2014.

[46] F. Pittala, I. Slim, A. Mezghani, and J. A. Nossek. Training-Aided Frequency-Domain Channel Estimation and Equalization for Single-Carrier Coherent Optical Transmission Systems. *Journal of Lightwave Technology*, 32(24) :4247–4261, December 2014.

[47] Vincent Poor. *An Introduction of Signal Detection and Estimation*. Springer, 1989.

[48] J. G. Proakis and D. G. Manolakis. *Digital signal processing : principles, algorithms, and applications*. Prentice Hall, 1996.

[49] M. Qiu, Q. Zhuge, X. Xu, M. Chagnon, M. Morsy Osman, and D. V. Plant. Simple and efficient frequency offset tracking and carrier phase recovery algorithms in single carrier transmission systems. *Optics Express*, 21(7) :8157–8165, April 2013.

[50] S. Savory. Digital Filters for coherent optical receivers. *Optics Express*, 16(2) :804, 2008.

[51] R. Schmogrow, P.C. Schindler, C. Koos, W. Freude, and J. Leuthold. Blind Polarization Demultiplexing With Low Computational Complexity. *IEEE Photonics Technology Letters*, 25(13) :1230–1233, January 2013.

[52] M. Seimetz. Laser linewidth limitations for optical systems with highorder modulation employing feed forward digital carrier phase estimation. In *Optical Fiber Communication Conference*, page OTuM2, March 2008.

[53] M. Selmi, Y. Jaouen, and P. Cibalt. Accurate digital frequency offset estimator for coherent PolMux QAM transmission systems. In *European Conference and Exhibition on Optical Communication*, page P3.08, Vienna, Austria, September 2009.

[54] M.-Z. Shao and N. Badler. Spherical sampling by Archimedes' theorem. Technical Report MS-CIS-96-02, University of Pennsylvania, Department of Computer and Information Science, 1996.

[55] K. Shi, R. Watts, D. Reid, T.N. Huynh, C. Browning, P.M. Anadarajah, F. Smyth, and L. P. Barry. Dynamic Linewidth Measurement Method via an Optical Quadrature Front End. *IEEE Photonics Technology Letters*, 23(21) :1591 – 1593, 2011.

[56] W. Shieh and K.-P. Ho. Equalization enhanced phase noise for coherent detection systems using electronic digital signal processing. *Optics Express*, 16(20) :15718–15727, 2008.

[57] R. A. Soriano, F. N. Hauske, N. G. Gonzalez, Z. Zhang, Y. Ye, and I. T.Monroy. Chromatic Dispersion Estimation in Digital Coherent Receivers. *IEEE Journal of Lightwave Technology*, 29(11) :1627 –1637, June 2011.

[58] J. Stanley. High-Speed ASIC for Optical Communications. In *Optical Fiber Communications Conference*, page Th3E.7, San Francisco, CA, USA, 2014.

[59] N. Stojanovi, B. Mao, and Y. Zhao. Digtial phase detector for nyquist and faster than nyquist systems. *IEEE Communications Letters*, 18(3) :511–514, March 2014.

[60] N. Stojanovi, Y. Zhao, and C. Xie. Feed-Forward and Feedback Timing Recovery for Nyquist and Fast than Nyquist Systems. In *Optical Fiber Communication Conference*, page Th3E.3, 2014.

[61] N. Stojanovic, C. Xie, Y. Zhao, B. Mao, and N. G. Gonzalez. A circuit enabling clock extraction in coherent receivers. In *European Conference and Exhibition on Optical Communication*, page P3.08, 2012.

[62] H. Sun and K.-T. Wu. A novel dispersion and PMD tolerant clock phase detector for coherent transmission. In *Optical Fiber Communication Conference*, page OMJ4, Los Angeles, CA, USA, 2011.

[63] B. Szafraniec, T. S. Marshall, and B. Nebendahl. Performance Monitoring and Measurement Techniques for Coherent Optical Systems. *Journal of Lightwave Technology*, 31(4) :648– 663,, February 2013.

[64] B. Szafraniec, B. Nebendahl, and T. Marshall. Polarization Demultiplexing in Stokes Space. *Optics Express*, 18(17) :17928–17939, August 2010.

[65] M. G. Taylor. Phase Estimation Methods for Optical Coherent Detection Using Digital Signal Processing. *Journal of Lightwave Technology*, 27(7) :901–914, 2009.

[66] F. Vacondio, C. Simonneau, A. Voicila, E. Dutisseuil, J.-M. Tanguy, J.-C. Antona, G. Charlet, and S. Bigo. Real time implementation of packet-by-packet polarization demultiplexing in a 28 Gb/s burst mode coherent receiver. In *Optical Fiber Communication Conference*, page OM3H.6, Los Angeles, California United States, March 2012.

[67] H. E. Weste and D. M. Harris. *CMOS VLSI design : a circuts and systems perspective*. Pearson Education, Inc., Addison Wesley, 2011.

[68] P. J. Winzer and G. J. Foschini. Mode Division Multiplexed Transmission System. In *Optical Fiber Communication Conference*, page Th1J.1, 2014.

[69] P. J. Winzer, A. H. Gnauck, C. R. Doerr, M. Magarini, and L. L. Buhl. Spectrally efficient long-haul optical networking using 112 Gbps polarization multiplexed 16 QAM. *IEEE Journal of Lightwave Technology*, 28(4) :547–556, Feb 2010.

[70] P. J. Winzer, A. H. Gnauck, A. Konczykowska, F. Jorge, and J. Y. Dupuy. Penalties from In Band Crosstalk for Advanced Optical Modulation Formats. In *European Conference and Exhibition on Optical Communication*, page Tu.5.B, 2011.

[71] K.-T. Wu and H. Sun. Frequency-Domain Clock Phase Detector for Nyquist WDM Systems. In *Optical Fiber Communications Conference*, page Th3E.2, San Francisco, CA, USA,, 2014.

[72] C. Xie and S. Chandrasekhar. Two-stage constant modulus algorithm equalizer for singularity free operation and optical performance monitoring in optical coherent receiver. *Optical Fiber Communication Conference*, page OMK3, March 2010.

[73] Z. Yu, X. Yi, Q. Yang, M. Luo, J. Zhang, L. Chen, and K. Qiu. Experimental Demonstration of Polarization Demultiplexing in Stokes Space for Coherent Optical OFDM. In *Optical Fiber Communication Conference*, page OW3B, March 2013.

[74] Z. Yu, X. Yi, J. Zhang, M. Deng, H. Zhang, and K. Qiu. Modified Constant Moduolus algorithm with Polarization Demultiplexing in Stokes Space in Optical Coherent Receiver. *IEEE Journal of Lightwave Technology*, 31(19) :3203 – 3209, 2013.

[75] S. Zhang, C. Yu, P. Y. Kam, and J. Chen. Parallel Implementation of Decision-Aided Maximum-Likelihood Phase Estimation in Coherent M-ary Phase-Shift Keying Systems. *IEEE Photonics Technology Letters*, 21(19) :1363–1365, October 2009.

[76] X. Zhou, L. E. Nelson, P. Magill, R. Isaac, B. Zhu, D. W. Peckham, P. I. Borel, and K. Carlson. High Spectral Efficiency 400 Gbps Transmission Using PDM Time Domain Hybrid 32/64 QAM and Training Assisted Carrier Recovery. *IEEE Journal of Lightwave Technology*, 31(7) :999–1005, April 2013.

[77] Q. Zhuge, M. E. Pasandi, X. Xu, B. Chatelain, Z. Pan, and M. Osman. Linewidth tolerant low complexity pilot aided carrier phase recovery for M QAM using superscalar parallelization. In *Optical Fiber Communication Conference*, page OTu2G.2, Los Angeles, CA, USA, 2012.