

Iterative Graphical Algorithms for Phase Noise Channels

Seare Haile Rezenom

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Electronic Engineering
University of KwaZulu-Natal
South Africa

As the supervisor of the candidate, I agree to the submission of this thesis.

Prof. Fambirai Takawira

Sign.

December 2020

Declaration

I declare that

- i) I am aware of the policy of the University on original work.
- ii) The material presented in this thesis is my own original work.
- iii) The relevant sources have been referenced explicitly.
- iv) This thesis has not been submitted towards obtaining any degree, or qualification, at any other university or tertiary institute.

Seare H. Rezenom

Sign.

Abstract

This thesis proposes algorithms based on graphical models to detect signals and characterise the performance of communication systems in the presence of Wiener phase noise. The algorithms exploit properties of phase noise and consequently use graphical models to develop low complexity approaches of signal detection. The contributions are presented in the form of papers.

The first paper investigates the effect of message scheduling on the performance of graphical algorithms. A serial message scheduling is proposed for Orthogonal Frequency Division Multiplexing (OFDM) systems in the presence of carrier frequency offset and phase noise. The algorithm is shown to have better convergence compared to non-serial scheduling algorithms.

The second paper introduces a concept referred to as circular random variables which is based on exploiting the properties of phase noise. An iterative algorithm is proposed to detect Low Density Parity Check (LDPC) codes in the presence of Wiener phase noise. The proposed algorithm is shown to have similar performance as existing algorithms with very low complexity.

The third paper extends the concept of circular variables to detect coherent optical OFDM signals in the presence of residual carrier frequency offset and Wiener phase noise. The proposed iterative algorithm shows a significant improvement in complexity compared to existing algorithms.

The fourth paper proposes two methods based on minimising the free energy function of graphical models. The first method combines the Belief Propagation (BP) and the Uniformly Re-weighted BP (URWBP) algorithms. The second method combines the Mean Field (MF) and the URWBP algorithms. The proposed methods are used to detect LDPC codes in Wiener phase noise channels. The proposed methods show good balance between complexity and performance compared to existing methods.

The last paper proposes parameter based computation of the information bounds of the Wiener phase noise channel. The proposed methods compute the information lower and upper bounds using parameters of the Gaussian probability density function. The results show that these methods achieve similar performance as existing methods with low complexity.

To my parents

Acknowledgement

I would like to express my profound gratitude to my supervisor, Prof. Fambirai Takawira, for his guidance and supervision of this project. I appreciate his ability to simplify a complex subject matter and subsequently formulate challenging questions. This approach to inquiry has stimulated the research work. I would like to thank Dr. Dieter Duyck for the interesting discussions on graphical models early on in this work. I also thank Prof. Jules Tapamo for the insightful discussions on complexity of algorithms.

I have no words that express my gratitude to all my family for the constant support and encouragement.

I also thank all my friends and colleagues for making this work seem bearable. Many thanks to the school administration assistant, Bennedin Mokoena, who has been helpful with administrative matters.

Contents

List of Acronyms	viii
List of Figures	ix
I Introduction	1
1 Wireless Communication Systems	2
2 Phase Noise	3
3 Orthogonal Frequency Division Multiplexing System	5
4 Graphical Models	6
5 Free Energy of Graphical Models	9
6 Motivation	14
7 Contributions	15
8 List of Publications	17
References	18
II Papers	24
1 Fast Detection of OFDM Systems Using Graphical Models	25
I Introduction	25
II System Model	27
III Message Construction	28
IV Results	29
V Conclusion	32
References	33
2 Iterative Low Complexity Algorithm for LDPC Systems in the Presence of Phase Noise	35
I Introduction	35
II Problem Formulation	36
III Proposed Method	38
IV Results	41
V Conclusion	43
Appendix	44
References	45

3	Iterative Detection of CO-OFDM Signals with Phase Errors Using Circular Random Variables	46
I	Introduction	46
II	System Model	48
III	Proposed Receiver	50
IV	Results	53
V	Conclusion	55
	References	56
4	Graphical Algorithms for Optical LDPC Systems with Phase Noise Errors	59
I	Introduction	59
II	Background	62
III	System Model	63
IV	Proposed Algorithms	64
V	Complexity	69
VI	Results	69
VII	Conclusion	72
	Appendix A	73
	Appendix B	75
	References	77
5	Parameter Based Computation of Information Bounds for Wiener Phase Noise Channels	80
I	Introduction	80
II	Problem Formulation	82
III	Proposed Lower Bound	82
IV	Proposed Upper Bound	84
V	Results	87
VI	Conclusion	89
	References	90
III	Conclusion	92
1	Conclusion	93
2	Future work	94
	References	96

List of Acronyms

AMP	Approximate Message Passing
AWGN	Additive White Gaussian Noise
BER	Bit Error Rate
BP	Belief Propagation
BPSK	Binary Phase Shift Keying
CDMA	Code Division Multiple Access
CFO	Carrier Frequency Offset
CP	Cyclic Prefix
DFT	Discrete Fourier Transform
DVB	Digital Video Broadcasting
EDGE	Extended Data rates for GSM Evolution
EHF	Extremely High Frequency
EP	Expectation Propagation
FBMC	Filter Bank Multi-Carrier
GAMP	Generalised Approximate Message Passing
GFDM	Generalised Frequency Division Multiplexing
GPRS	Generalised Packet Radio Service
GSM	Global System for Mobile communication
IDFT	Inverse Discrete Fourier Transform
LDPC	Low Density Parity Check
MF	Mean Field
MIMO	Multiple Input Multiple Output
OFDM	Orthogonal Frequency Division Multiplexing
PN	Phase Noise
PSK	Phase Shift Keying
QAM	Quadrature Amplitude Modulation
QPSK	Quaternary Phase Shift Keying
RF	Radio Frequency
SNR	Signal to Noise Ratio
UFMC	Universal Filter Multi-Carrier
URWBP	Uniformly Reweighted Belief Propagation

List of Figures

Introduction

Figure 1	A basic communication system.	3
Figure 2	Graphical model.	7
Figure 3	Factor Graph.	11
Figure 4	Region based graphical representation.	12

Paper 1

Figure 1	Graphical model.	28
Figure 2	Convergence performance of different schedules at 30 dB.	31
Figure 3	Error rate performance versus SNR for the different schedules.	32

Paper 2

Figure 1	The graphical model showing the code graph and the phase noise graph.	38
Figure 2	Bit error rate performance for BPSK and QPSK.	42
Figure 3	Bit error rate performance for 8-PSK.	43

Paper 3

Figure 1	Coherent optical OFDM system model.	49
Figure 2	Graphical model.	50
Figure 3	Estimation of phase errors with the graphical models.	53
Figure 4	Error rate performance of the proposed algorithm.	54

Paper 4

Figure 1	Basic description of a coded system.	62
Figure 2	Description of a code-based transmitter.	63
Figure 3	Systemic description of graph-based algorithm.	64
Figure 4	Graphical model of (2).	64
Figure 5	BER of the proposed algorithms for varying ρ	71
Figure 6	BER of the proposed algorithms.	71
Figure 7	BER as a variation of the normalised laser linewidth.	72

Paper 5

Figure 1	Information bounds of the proposed methods 4-QAM.	88
Figure 2	Information bounds of the proposed methods 16-QAM.	88

I

Introduction

Introduction

1 Wireless Communication Systems

The technology of wireless communication systems has shown significant advances over the past decades. The First Generation (1G) wireless communication systems were capable of offering voice services only. The ability to transmit voice and data digitally was possible in the Second Generation (2G) wireless communications systems which use digital modulation techniques to offer voice and limited data services. This was implemented using technologies based on the Global System for Mobile communication (GSM) and the Code Division Multiple Access (CDMA) system. Further advances of these technologies have been achieved using the Generalised Packet Radio Services (GPRS), Extended Data rates for GSM Evolution (EDGE), etc. The Third Generation (3G) wireless communication systems have been characterised by a significant improvement of data rates compared to the 2G systems, and use the wide-band CDMA system as the underlying technology. The Fourth Generation (4G) wireless communication systems have increased the data rates using the Orthogonal Frequency Division Multiplexing (OFDM) system, and are further evolving by including many techniques such as the multiple antennas and carrier aggregation [1], [2]. The Fifth Generation (5G) wireless communication systems aim to achieve very low latency, high connectivity, and very high data rates. These objectives can be achieved with the use of enhanced multiple carrier techniques [3]–[5], massive MIMO (multiple-input multiple-output) [6], millimetre wave (mmWave) frequency transmissions [7], and other technologies [8]. Beyond 5G wireless communication systems are likely to be based on the non-orthogonal multiplexing access [9] and other emerging technologies that counteract the effect of the channel using smart reflecting surfaces [10], [11].

A basic wireless communication system is shown in Figure 1. The binary information signal is encoded to counteract errors introduced by the wireless channel. The coded information bits are then mapped using symbols selected from a predefined set. The mapped sequences are further processed with appropriate transmitter techniques followed by frequency up-conversion. The transmitted Radio Frequency (RF) signal encounters the effects of a wireless channel. The received signal, which at this point includes the effects of the Additive White Gaussian Noise (AWGN), is processed using inverse operations such as frequency down-conversion, analogue-to-digital conversion, demapping and decoding to recover the transmitted information bits. The received sequence is also affected by device

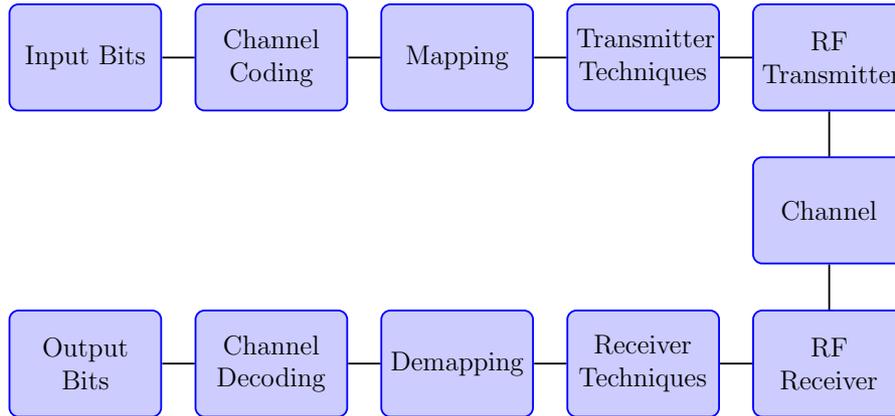


Figure 1: A basic communication system.

uncertainties such as the carrier frequency offset, phase noise, etc. Hence, the received signal has to be processed further using specialised algorithms to recover the transmitted bits in the presence of these uncertainties. Therefore, the emphasis of this thesis is on the design of receiver algorithms that counteract the effect of device uncertainties, particularly that of the phase noise.

This chapter is organised as follows. Section 2 discusses the oscillator phase noise. In Section 3, a brief description of an OFDM system is presented. In Section 4, the analysis tools used in this thesis are discussed. In Section 5, the minimisation method of the free energy function is presented. Section 6 presents the motivation of the methods proposed in this work. Section 7 lists the contributions of this thesis. Finally, Section 8 presents the list of publications.

2 Phase Noise

An ideal oscillator generates a waveform which has a stable frequency. In a practical oscillator however, the generated waveform is not stable. This is due to the inherent characteristics of oscillators which causes the period of oscillation to drift with time. The time-domain waveform of an ideal oscillator can be written as $y(t) = Ax(\omega_0 t)$, where A is the amplitude and $x(\cdot)$ represents a periodic waveform. In contrast, the waveform of a practical oscillator can be described as $y(t) = (1 + A(t))x(\omega_0 t + \phi(t))$ where $\phi(t)$ describes the drift in time and $A(t)$ is the amplitude variation. In oscillatory systems, the amplitude variation is not a significant source of phase noise [12],[13] and its effect is mostly ignored. The frequency spectrum of an ideal oscillator is modelled as an impulse signal centred at the frequency of oscillation. The frequency spectrum of a practical oscillator is represented by an envelop that spreads around the carrier frequency. This frequency spectrum is commonly referred to as phase noise. The nature of this spectrum depends on the quality of the oscillator and the quantifying parameters of oscillator phase noise are derived from it.

There are various noise sources that affect the frequency characteristics of an oscillator. These sources have been identified as the thermal noise, flicker noise, shot noise, substrate and power-supply side noise etc., and their impact on the frequency spectrum of an oscillator has been studied extensively [12], [14]. Based on the impact of the noise sources, a mathematical model for the phase noise can be developed. For a general application, the phase noise has been modelled as a Wiener process, and this modelling is valid if the oscillator is assumed to be affected by thermal noise only [15]. Wiener phase noise has been used by many authors in the general study of communication systems [16]–[19]. For specific applications that require more accuracy, the effect of other noise sources has to be considered. A recent study [15] provides the use of a more accurate phase noise model in communication systems.

It may be necessary to define metrics so that the designer may quantify and compare practical oscillators. One such metric that is widely applied uses the ratio of the single side-band power measured at a predefined offset from the oscillator carrier frequency to the power of the carrier frequency [13]. This ratio may be described as $L(\Delta f) = \frac{P_{\Delta f}}{P_{centre}}$, where Δf represents the predefined frequency offset from the oscillator centre frequency. Furthermore, $P_{\Delta f}$ represents the power at the predefined offset, and P_{centre} is the power of the carrier frequency. This ratio is described in units of dBc/Hz.

Another metric is obtained by analysing the time-domain dynamics of oscillators. Similar to the drift in the frequency domain, the period of a noisy oscillator drifts as a function of time. This implies that the period of the waveform is different for each cycle. The statistical variance of these timing variations, commonly referred to as timing jitter, is widely used in modelling practical oscillators in time-domain. In fact, the variance can be obtained by analysing the frequency spectrum of the oscillator. The reader may refer to the work in [20], [21] for more details. The frequency characteristics of a practical oscillator may be used to explain the mechanism by which the performance degrades as the oscillator is used in communication systems—which uses local oscillator for the up-conversion and down-conversion of the base-band signal.

Phase noise reduces the available signal-to-noise ratio of the communication system. For single carrier and multi-carrier systems, it has been shown that this reduction due to phase noise is described as $10 \log(1 + \sigma_\phi^2 \gamma)$, where σ_ϕ^2 refers to the variance of the phase noise and γ refers to the signal-to-noise ratio [22]. In multi-carrier systems, the shift from the desired oscillator frequency leads to interference from nearby carriers.

In this thesis, the phase noise of an oscillator is modelled as a Wiener process. This requires successive phase noise samples to be computed as

$$\phi(t) = \phi(t - \tau) + \Delta\phi, \quad (1)$$

where τ is the measure of the time delay, and $\Delta\phi$ is a Gaussian random variable with zero

mean and variance σ_ϕ^2 . Therefore, a transmitted signal $x(t)$ is received as

$$r(t) = x(t) e^{j\phi(t)} + w(t), \quad (2)$$

where $w(t)$ is an additive white Gaussian noise with zero mean and variance σ^2 .

3 Orthogonal Frequency Division Multiplexing System

The use of multiple carriers in a communication system increases the available data rate when compared to single-carrier systems [23]. In multi-carrier systems, the information signals are transmitted using different frequencies and this increase the available data rate. Furthermore, since the transmitted signal is spread across multiple frequencies, referred to as subcarrier frequencies, the system handles the frequency-selective fading channel better than single carrier systems. If the spacing between successive subcarrier frequencies is selected to be smaller than the coherence bandwidth of the channel, then each subcarrier frequency would only experience flat fading. These two properties have made multiple carrier systems very attractive in communication systems.

In an OFDM system, the information signal $\mathbf{a} = [a_1, a_2, \dots, a_N]$ is spread across N subcarrier frequencies. An Inverse Discrete Fourier Transform (IDFT) operation is used to generate the transmitted signal. In order to eliminate the effect of the multipath channel which results in successive symbols to arrive simultaneously, the output of the IDFT is extended by appending data sequences whose length is longer than the delay spread of the channel. The sequence appended is referred to as guard interval and is widely implemented in two ways. In first case, the last part of the output sequence of the IDFT is appended to the first part of the output sequence. This is called the cyclic prefix method [23]. In the second case, an all zeros sequence is appended to the first part of the signal from the output of the IDFT. This is referred to as zero padding method [24]. In each case, the objective is to prevent an inter-symbol interference in the received signal. Other guard interval methods are discussed in [25],[26] while a comparison of these methods is discussed in [27].

The time domain waveform at the output of the IDFT is described by

$$x(t) = \sum_{k=0}^{N-1} a_k e^{j2\pi kt/T}, \quad (3)$$

where a_k is the information symbol that is fed to the IDFT and T is the period of the OFDM symbol. A cyclic prefix is then appended to $x(t)$ to form the signal $x_a(t)$. The received signal $r(t)$ is described as

$$r(t) = x_a(t) * h(t) + w(t), \quad (4)$$

where $h(t)$ is the multipath channel, $*$ denotes the convolution operator and $w(t)$ is the

additive White Gaussian noise. After filtering and sampling, the discrete received signal \mathbf{r} is fed to a Discrete Fourier Transform (DFT), which results in $\mathbf{y} = \mathbb{F}\{\mathbf{r}\}$. where \mathbb{F} is the DFT operator with entries $\frac{1}{\sqrt{N}} \exp(-j2\pi pq/N)$ where $p, q \in [0, N - 1]$.

If one considers a communication system in the presence of carrier frequency offset and phase noise, the transmitted signal can be written as

$$x(t) = e^{j2\pi vt/T} e^{j\phi(t)} \sum_{k=0}^{N-1} a_k e^{j2\pi kt/T}, \quad (5)$$

where v represents the carrier frequency offset normalised to the subcarrier spacing and $\phi(t)$ is the phase noise. It has been investigated by many authors that the presence of carrier frequency offset and phase noise seriously degrades the performance of OFDM systems, and the estimation of these two parameters is crucial in order to realise the advantages of the OFDM system [16], [18], [28], [29].

4 Graphical Models

The system model of a general communication system can be written in the form of $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}$ where \mathbf{y} is the received signal, \mathbf{A} represents additional information about the system, \mathbf{x} is the information signal and \mathbf{w} is the additive white Gaussian noise. The objective may differ depending on what information is available. One objective would be to find the transmitted signal \mathbf{x} assuming \mathbf{A} is known. Another objective would be to find \mathbf{A} assuming the receiver has knowledge of the information signal in the form of pilots or other known sequences. The objective could also be to jointly detect the system information \mathbf{A} and the information signal \mathbf{x} . Although the objective varies, the problem is posed as an inverse mathematical problem and many techniques can be employed to find the unknown parameters of the system.

In this thesis, the emphasis is on the use of probabilistic approaches to solve the inverse problem. For some problems, the global probability density function of the system can be factorised into local probability density functions. This allows one to develop a visual framework that reveals the interaction of the parameters within the system. This visual framework is referred to as graphical model or factor graph. A good review of graphical models is presented in [30], [31].

To describe a given problem using graphical models, consider a problem of N variables with a joint probability density function $f(x_1, x_2, \dots, x_N)$. The following example describes the formulation of a graphical model from the global probability density function. Assume the global probability density function $f(x_1, x_2, \dots, x_N)$ can be factorised into localised probability density functions $\prod_i f_i(\mathbf{x}_i)$, where \mathbf{x}_i represents all variables of the local distribution function f_i . Therefore, a model can be developed that shows a graphical relationship between the variable nodes. For example, if the joint probability density

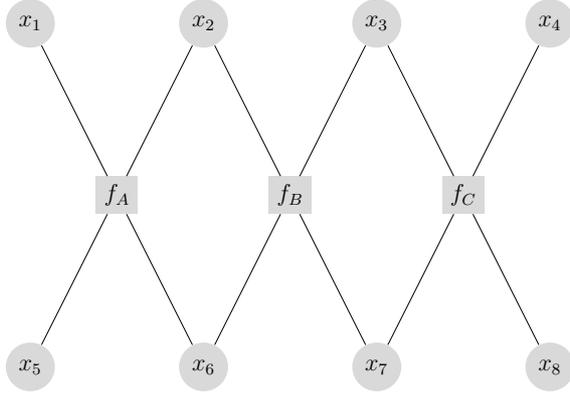


Figure 2: Graphical model.

function of a problem can be described as

$$f(x_1, x_2, \dots, x_8) = f_A(x_1, x_2, x_5, x_6) f_B(x_2, x_3, x_6, x_7) f_C(x_3, x_4, x_7, x_8). \quad (6)$$

This probability density function is described using a graph and is shown in Figure 2. In this case, f_A , f_B and f_C are the local probability density functions.

The variables x_1, x_2, \dots, x_N are referred to *variable nodes* and the local probability density functions $f_A(\cdot)$, $f_B(\cdot)$ and $f_C(\cdot)$ are referred to as *factor nodes*. The resulting graphical model \mathcal{G} can be represented mathematically as $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{F})$, where \mathcal{V} represents the set of variable nodes, \mathcal{E} represents the set of edges, and \mathcal{F} represents the set of factor nodes.

Once a given problem is represented using a graphical model, each variable node and factor node is made to construct and propagate messages to nearby nodes. This process is repeated for a defined number of iterations, or until the messages propagated from each node show negligible variations. Such a method of propagating messages in a graphical model has been proposed in [32] and thereafter has been widely used. The messages that propagate in the graphical model are usually constructed using a method referred to as the Belief Propagation (BP), or the sum-product algorithm.

For a given graphical model, the BP algorithm requires the computation of messages from each variable node and the factor node. The message from a factor node a to a variable node i is constructed as

$$m_{a \rightarrow i}(x_i) = \sum_{\mathbf{x}_a \sim x_i} f(\mathbf{x}_a) \prod_{j \in \mathcal{N}(a) \sim i} m_{j \rightarrow a}(x_j), \quad (7)$$

where \mathbf{x}_a is used to denote the set of variable nodes connected to the factor node a , and $\mathbf{x}_a \sim x_i$ represents the set of variable nodes connected to factor node a with the exception of variable node x_i . Further, $\mathcal{N}(a)$ is used to denote any neighbouring nodes of factor node a .

The message from a variable node i to a factor node a is constructed as

$$m_{i \rightarrow a}(x_i) = \prod_{b \in \mathcal{N}(i) \sim a} m_{b \rightarrow i}(x_i). \quad (8)$$

When all the messages are computed, the marginal belief $b_i(x_i)$ of each node x_i is computed as

$$b_i(x_i) = \prod_{a \in \mathcal{N}(i)} m_{a \rightarrow i}(x_i). \quad (9)$$

There are other non-BP algorithms of constructing messages proposed in the existing work [33]–[36]. These non-BP algorithms have their respective advantages and disadvantages depending on the type of problem. For problems which are represented by discrete variables, the belief propagation algorithm has been widely used with good performance. For problems with continuous variables, non-BP based methods have been shown to provide good performance [36].

The principle behind the good performance of the belief propagation algorithm has been difficult to explain until the authors in [33] showed an interesting insight when they formulated it as a constrained optimisation problem. The objective function has been selected as the free energy of the graphical model. In order to determine the free energy of the graphical model, the authors used the principle used in statistical physics, where the probability of a state varies inversely with its energy. By using this analogy, the objective function has been optimised under some constraints. The insight derived from this method shows that the solution to the optimisation problem is the belief propagation algorithm. The resulting message propagation algorithm that optimises the objective function may change depending on the structure of the graph and the set of constraints imposed on the graph. This insight has proved to be very crucial to explain the underlying principle of the belief propagation methods and to further develop algorithms depending on the actual graphical model of the system.

Despite the established relationship between the free energy function and the belief propagation algorithm, the convergence of belief propagation continues to be a challenge. It has been widely established that belief propagation algorithm achieves good performance when it converges. However, there is no guarantee that the belief propagation algorithm always converges. It has also been shown that the belief propagation algorithm may diverge, or oscillate between some values [37], [38]. In order to improve the convergence properties of the belief propagation algorithm, message damping has been shown to reduce message divergence or message oscillation [39]. However, the convergence of the belief propagation algorithm still continues to be a challenge.

5 Free Energy of Graphical Models

Background

In this section, we present an overview of the formulation of the free energy function of graphical models. We use $\mathcal{G} = (\mathcal{V}, \mathcal{F}, \mathcal{E})$ to refer to a graphical structure \mathcal{G} consisting of the set of \mathcal{V} variable nodes, the set of \mathcal{F} factor nodes, and the set of \mathcal{E} edges. The graphical model $\mathcal{M} = (\mathcal{G}, \mathbf{f})$ is used to refer to a graph \mathcal{G} , whose factor nodes subscribe to a set of functions f_a , i.e., $\{f_a : a \in \mathcal{F}\}$. Further notations used are described in the table below.

Table 1: Description of Symbols.

Symbol	Description
i	variable node index
a	factor node index
(ai)	an edge connecting factor node a to variable node i
G_a	a Bethe region [33]
\mathbf{x}_{G_a}	variable nodes in the graph G_a
\mathbf{x}_a	a set of variable nodes connected to the factor f_a
\mathcal{V}	variable nodes of a graph
$b_i(x_i)$	belief of variable node i
$b_{G_a}(\mathbf{x}_{G_a})$	belief of the graph G_a
$\mathcal{N}(i)$	neighbourhood of a variable node i
$\mathcal{N}(a)$	neighbourhood of factor node a
\sim	excluding

Free Energy

Consider the joint probability density function $f_{X_0 X_1 \dots X_{N-1}}(\mathbf{x})$ of a multi-variable \mathbf{x} . Using the theoretical framework of the free energy, there can be defined an energy function for this probability density function taking into account the average energy of the graph and the total measure of uncertainty – entropy of the graph. The concept of free energy aims to determine the state $\mathbf{x}^* = \{x_0^*, x_1^*, \dots, x_{N-1}^*\}$ that minimises the total *free energy* content of the graph describing the factorisation of the density function.

The energy content of a state \mathbf{x} is defined as [33]

$$E(\mathbf{x}) = - \sum_{a=1}^M \log f_a(\mathbf{x}_a). \quad (10)$$

Assume a trial probability density function $b(\mathbf{x})$ can be obtained for the true joint density function $p(\mathbf{x})$, then the free energy of the approximating density function is given by [33]

$$F(b) = U(b) - H(b), \quad (11)$$

where $U(b)$ is the average energy of the function $b(\mathbf{x})$ and is given by

$$U(b) = \sum_{\mathbf{x}} b(\mathbf{x})E(\mathbf{x}), \quad (12)$$

and $H(b)$ is the entropy of the function $b(\mathbf{x})$ defined as

$$H(b) = - \sum_{\mathbf{x}} b(\mathbf{x}) \log b(\mathbf{x}). \quad (13)$$

Region Based Free Energy

There have been recent efforts that apply the region-based free energy minimisation in graphical models which was originally observed by [33]. The region-based free energy is a technique where the free energy of a graph \mathcal{G} is estimated by sub-dividing the graph into many regions. The free energies of the sub-divided R regions are computed separately using standard methods, and finally the individual computation of the regions are combined by accounting for any overlap among regions.

The free energy of the sub-divided region is given by

$$F_R(b) = U_R(b_R) - H_R(b_R), \quad (14)$$

where $U_R(b_R)$ is the average region energy of the function $b_R(\mathbf{x}_R)$ and is given by

$$U_R(b_R) = \sum_{\mathbf{x}_R} b_R(\mathbf{x}_R)E_R(\mathbf{x}_R), \quad (15)$$

and $H_R(b_R)$ is the region entropy of the function $b_R(\mathbf{x}_R)$ defined as

$$H_R(b_R) = - \sum_{\mathbf{x}_R} b_R(\mathbf{x}_R) \log b_R(\mathbf{x}_R). \quad (16)$$

An important step in this process is identifying the criteria for region selection. Although regions can be selected in many ways, as are stated in [33], a particular method of region selection called the *Bethe* method is presented. In the Bethe method [33], regions are classified as small and large regions. Small regions contain single variable nodes only while a large region contains a single factor node and all the variable nodes connected to it.

The region based free energy of the graph \mathcal{G} can be determined by

$$F_{\mathcal{G}}(b) = U_{\mathcal{G}}(\{b_{G_a}\}) - H_{\mathcal{G}}(\{b_{G_a}\}), \quad (17)$$

where the region based average energy is given by

$$U_{\mathcal{G}}(b) = \sum_{G_a \in \mathcal{G}} c_{G_a} U_{G_a}(b_{G_a}), \quad (18)$$

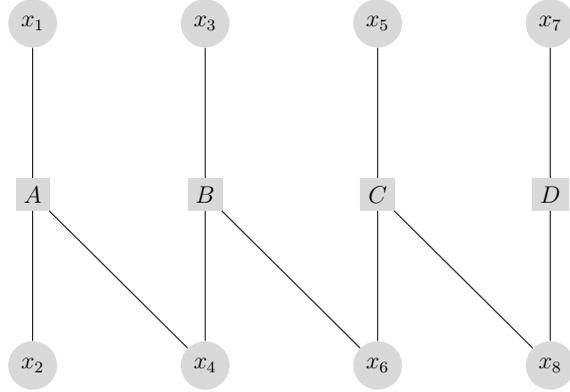


Figure 3: Factor Graph.

and the region based entropy is given by

$$H_G(b) = \sum_{G_a \in \mathcal{G}} c_{G_a} H_{G_a}(b_{G_a}), \quad (19)$$

where c_{G_a} is the counting number of each region. For the Bethe method of small region selection, the counting number is related to the degree of node c_i of a variable node i as $c_{G_a} = c_i$. The large Bethe region has a counting number of $c_{G_a} = 1$.

Example

In this example we show a derivation of the free energy of the graphical model described by the distribution

$$\begin{aligned} P_{X_1 \dots X_8}(x_1, \dots, x_8) &= \frac{1}{Z} \prod_{a=A}^D f_a(\mathbf{x}_a) \\ &= \frac{1}{Z} f_A(x_1, x_2, x_4) f_B(x_3, x_4, x_6) f_C(x_5, x_6, x_8) f_D(x_7, x_8), \end{aligned} \quad (20)$$

which consist of variable nodes $\mathcal{V} = \{x_1, \dots, x_8\}$, factor nodes $\mathcal{F} = \{A, B, C, D\}$ and edges $\mathcal{E} = \{(1A), (2A), (4A), (3B), (4B), (6B), (5C), (6C), (8C), (7D), (8D)\}$. Further, Z is a normalisation constant. The graphical model \mathcal{M} of the graph $\mathcal{G} = (\mathcal{V}, \mathcal{F}, \mathcal{E})$ can be denoted as $\mathcal{M} = (\mathcal{G}, \mathbf{f})$, where \mathcal{G} shows the graph described and $\mathbf{f} = f_a : a \in \mathcal{F}$. Figure 3 shows the graphical model of this distribution.

The Bethe method of region selection is shown in Figure 4, which consist of small and large regions. The shaded regions represent the large region selection. The small regions are represented by the respective variable nodes. The counting number of each large region is 1. The counting numbers of the small region varies depending on the number of overlaps. For example, the variable nodes x_4 , x_6 and x_8 have a counting number of 2. The remaining variable nodes have a counting number of 1.

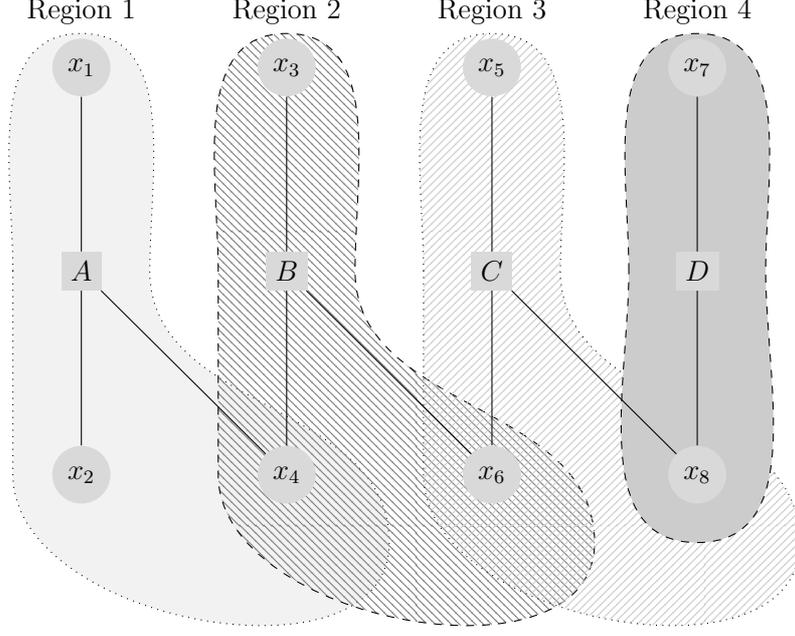


Figure 4: Region based graphical representation.

The average energy $U_{\mathcal{G}}(\{b_{\mathcal{G}}\})$ of the large region is

$$\begin{aligned}
 U_{\mathcal{G}}(b) &= \sum_{G_a \in \mathcal{G}} c_{G_a} U_{G_a}(b_{G_a}) \\
 &= -b_{G_A}(\mathbf{x}_{G_A}) \log f_{G_A}(\mathbf{x}_{G_A}) - b_{G_B}(\mathbf{x}_{G_B}) \log f_{G_B}(\mathbf{x}_{G_B}) - b_{G_C}(\mathbf{x}_{G_C}) \log f_{G_C}(\mathbf{x}_{G_C}) \\
 &\quad - b_{G_D}(\mathbf{x}_{G_D}) \log f_{G_D}(\mathbf{x}_{G_D}).
 \end{aligned} \tag{21}$$

It can be seen that the counting number c_{G_a} of each large region is 1, i.e., $c_{G_a} = 1, \forall G_a \in \mathcal{G}$. Similarly the entropy of the large region is determined as

$$\begin{aligned}
 H_{\mathcal{G}}(b) &= \sum_{G_a \in \mathcal{G}} c_{G_a} H_{G_a}(b_{G_a}) \\
 &= -b_{G_A}(\mathbf{x}_{G_A}) \log b_{G_A}(\mathbf{x}_{G_A}) - b_{G_B}(\mathbf{x}_{G_B}) \log b_{G_B}(\mathbf{x}_{G_B}) - b_{G_C}(\mathbf{x}_{G_C}) \log b_{G_C}(\mathbf{x}_{G_C}) \\
 &\quad - b_{G_D}(\mathbf{x}_{G_D}) \log b_{G_D}(\mathbf{x}_{G_D}).
 \end{aligned} \tag{22}$$

Finally, the entropy due to the small regions is determined as

$$\begin{aligned}
 H_{\mathcal{V}}(b) &= \sum_{i \in \mathcal{V}} H_i(b_i) - \sum_{i \in \mathcal{V}} c_i H_i(b_i) \\
 &= - \sum_{i \in \mathcal{V}} \sum_{x_i} (c_i - 1) b_i(x_i) \log(b_i(x_i)).
 \end{aligned} \tag{23}$$

Since the small Bethe region only consists of one variable node, it is valid to represent the resulting region as the set of variable nodes of the graph. It can be noted that when accounting for the entropy of the small regions, the number of overlapping regions should be considered and their effect removed from the total free energy. Hence, the variable nodes

that cause redundancy are the ones with a counting number greater than 1. Therefore, the free energy can be found by combining the previous three expressions as

$$F_G(b) = - \sum_{a=A}^D b_{G_a}(\mathbf{x}_{G_a}) \log f_{G_a}(\mathbf{x}_{G_a}) + \sum_{a=A}^D b_{G_a}(\mathbf{x}_{G_a}) \log b_{G_a}(\mathbf{x}_{G_a}) \quad (24)$$

$$+ \sum_{i \in \mathcal{V}} \sum_{x_i} (1 - c_i) b_i(x_i) \log b_i(x_i).$$

Fixed Points and Stationary Points

An iterative algorithm \mathcal{A} reaches convergence at a point or set of points known as *fixed points*. At this set of fixed points, or at the state $\mathbf{x}^* = \{x_0^*, x_1^*, \dots, x_{N-1}^*\}$, the Kullback-Leibler ¹ measure of divergence $\text{KL}(b(\mathbf{x}^*)||p(\mathbf{x}^*))$ is minimised. If this divergence measure is zero, the algorithm converges and the fixed points give the exact solution. Otherwise, the algorithm converges to an approximate solution or even diverges. Previous studies [33],[40] have shown the relationship between the BP fixed points and stationary points of the constrained Bethe free energy. The finding of the study is that the stationary points of the constrained Bethe based free energy are also the belief propagation fixed point [33]. The authors in [40] further highlighted that stable fixed points of belief propagation are the minima of the constrained Bethe free energy function.

Optimisation of the Free Energy Function

We can write the general formulation of the optimisation problem of an objective function $f(\mathbf{x})$ subject to a given set of constraints as [41]

$$\begin{aligned} & \text{minimise } f(\mathbf{x}), \mathbf{x} \in \mathbf{R}^n \\ & \text{subject to } g_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \\ & \quad \quad \quad h_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, p \end{aligned} \quad (25)$$

where $g_i(\mathbf{x})$ is the equality constraint function, $h_i(\mathbf{x})$ represents the inequality constraint function and \mathbf{R} is the set of real numbers. This problem can generally be solved using the Lagrangian method by minimising the function defined as a weighted sum of the objective function and the set of constraint functions as [41]

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\rho}) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^p \rho_j h_j(\mathbf{x}), \quad (26)$$

where λ_i and ρ_j are Lagrangian multipliers that correspond to the respective constraint functions, and $\{\boldsymbol{\lambda}, \boldsymbol{\rho}\}$ represents the collection of such multipliers.

¹The KL divergence measures the similarity between two distributions $b(x)$ and $p(x)$ by computing the metric $KL(b||p) = \int b(x) \log \frac{b(x)}{p(x)} dx$.

6 Motivation

In the previous sections, the construction of messages in the graphical model has been discussed. Once the messages are constructed, it is necessary to define a system to propagate these messages which is referred to as message scheduling. There has been various message scheduling methods in the published work. In the context of LDPC codes, one of the earliest is the so called flooding schedule, where the factor nodes propagate messages to the variable nodes, and subsequently the variable nodes propagate messages to the factor nodes [30]. Other scheduling methods serialise the propagation of messages using either the factor nodes or the variables nodes as a reference [42]–[49]. In a broader view, the problem of message scheduling is concerned with finding a method that activates the edges of a graphical model in a particular sequence. In [50], a scheduling method referred to as residual belief propagation is proposed where the nodes are activated according to a cost function. This approach is extended to the LDPC system in [51] and is referred to as informed dynamic scheduling. In effect, the main aspect of these approaches is to define metrics to activate a particular factor node or variable node. A recent study summarises various scheduling methods for the LDPC system [52]. It should be stated that message scheduling is associated with the complexity of the graphical model. In contrast to the LDPC graph, other systems may require graphical models with continuous variable nodes. The effect of scheduling for such graphical models is not extensively studied. This has motivated the work presented in Paper 1, where the OFDM system is considered in the presence of carrier frequency offset and phase noise, which are continuous random variables.

One of the problems of applying graphical models to solve communication problems is that the belief propagation algorithm usually results in intractable integrals. In the case of graphical models where the variables are discrete, one may compute and propagate probabilities at these discrete samples. One example of this is the decoding of binary LDPC codes where the variables constitute of two discrete samples. In fact using the logarithmic technique, the message at an arbitrary variable node is effectively represented using a single probability value at each sample. However, for graphical models where the variables are continuous, the use of discrete samples to represent continuous variables leads to high implementation complexity. There have been studies that reduce this complexity by modelling the messages using known distributions such as Gaussian density function [53], Tikhonov density function [19] and other parametric distributions. The use of these distribution functions have been shown to simplify the computation of the belief propagation algorithm. This difficulty of representing messages for continuous variables has motivated the work presented in Paper 2 which considers the detection of LDPC systems in the presence of phase noise. The circular properties of the phase noise are exploited to rewrite the sum-product algorithms to arrive at an efficient and low complexity algorithm for the detection of LDPC codes in the presence of phase noise. Furthermore, the work presented in Paper 3 considers the detection of coherent optical OFDM system in the presence of

residual carrier frequency offset and phase noise. The circular properties of these phase errors is exploited to derive tractable low complexity messages of the BP algorithm.

The graphical algorithms are increasingly used to model many existing and new problems in communication systems. One of the earliest uses of the belief propagation algorithm in communication systems has been in the decoding of capacity-approaching codes [54], [55]. The results obtained demonstrated that graphical models can be applied to a wide range of problems. Therefore, many authors have used graphical models to study problems in communication systems such as channel estimation [56], [57], frequency offset and phase noise estimation [58]–[60], cooperative localisation [61], spectrum sensing [62], [63], compressed sensing [64], [65], etc. The traditional algorithm used was the BP algorithm. However, this has advanced to include algorithms such as the Mean Field (MF) [34], Expectation Propagation (EP) [35], Approximate Message Passing (AMP) [64], Generalised Approximate Message Passing (GAMP) [66], Uniformly Re-weighted Belief Propagation (URWBP) [67], etc. There have been research efforts that combine two or more of these graphical algorithms such as the combined MF and BP [36]. This has motivated the work presented in Paper 4, where two graphical algorithms have been proposed using a combination of the BP, MF and URWBP algorithms. The performance of these algorithms have been characterised in LDPC systems in the presence of Wiener phase noise.

The methods of computing the mutual information bounds of communication systems over the additive white Gaussian noise channel is widely established in the published work. The performance bounds of this channel are also known for coded and uncoded systems. However, computing the mutual information limits for communication systems in the presence of phase noise, and in particular that of the Wiener phase noise, is a challenging problem. The theoretical bounds of the Wiener phase noise channel are not as easily computed as that of the additive white Gaussian noise channel [68], [69]. Therefore, the work presented in Paper 5 has been motivated by efforts to develop efficient algorithms to compute the information bounds of systems in the presence of Wiener phase noise.

7 Contributions

The contributions of this thesis are summarised as follows.

In Paper 1, various message scheduling methods are proposed for the detection of an OFDM system in the presence of carrier frequency offset and Wiener phase noise using graphical models. The proposed algorithms use serial and non-serial scheduling methods. The finding indicates the serial message scheduling converges faster than non-serial message scheduling methods.

In Paper 2, a low complexity algorithm is proposed to detect LDPC codes in strong Wiener phase noise channels. This method is unique since it models the phase noise variable nodes in the graph as circular random variables. This is shown to simplify the computation of the messages in the graphical model. The findings show that the proposed

algorithm has similar performance as the lowest complexity algorithm in the published work [19] while it reduces the complexity even further.

In Paper 3, an algorithm based on circular random variables is proposed and its performance is studied to detect coherent optical OFDM signals in the presence of residual carrier frequency offset and Wiener phase noise. The findings show that the proposed algorithm has significant improvement in performance and complexity when compared to algorithms that consider the discrete representation of phase noise [60].

In Paper 4, the concept of free energy minimisation is used to propose two graphical algorithms that combine the BP, MF and URWBP algorithms. The first algorithm combines the BP and URWBP graphical models. The second algorithm combines the MF and URWBP graphical models. The proposed algorithms have been applied to the LDPC detection problem in the presence of Wiener phase noise. The proposed algorithms represent the phase noise as a circular random variable, and an optimal reweighting factor is obtained which improves the performance of the non-optimised algorithms [70].

In Paper 5, low complexity methods are proposed to compute the mutual information bounds for communication systems in the presence of Wiener phase noise. The methods compute the lower and upper bounds of the mutual information using the parameters of the Gaussian probability density function, which have been used to approximate the phase noise distribution. The findings indicate that the proposed methods are effective for strong phase noise and for different modulation schemes. Further, the proposed approach has complexity advantages compared to existing algorithms [68], [69], [71].

8 List of Publications

The following list includes parts of the thesis that have been published or submitted for publication.

- (1) S. H. Rezenom and F. Takawira, “Fast Detection of OFDM Systems using Graphical Models,” in *Proceedings of IEEE Signal Processing Workshop*, pp. 154–158, UT, USA, 2015.
- (2) S. H. Rezenom and F. Takawira, “Iterative Low Complexity Algorithm for LDPC Systems in the Presence of Phase Noise,” *IEEE Wireless Communications Letters*, no. 6, pp. 794–797, 2017.
- (3) S. H. Rezenom and F. Takawira, “Iterative Detection of CO-OFDM Signals with Phase Errors Using Circular Random Variables,” *International Journal of Communication Systems (under review)*
- (4) S. H. Rezenom and F. Takawira, “Graphical Algorithms for Optical LDPC Systems with Phase Noise Errors,” *Physical Communications (to be submitted)*
- (5) S. H. Rezenom, “Parameter Based Computation of Information Bounds for Wiener Phase Noise Channels,” *IEEE Communications Letters (to be submitted)*

References

- [1] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, “Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications,” *IEEE Communications Magazine*, vol. 50, no. 2, 2012.
- [2] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*. Academic press, 2013.
- [3] B. Farhang-Boroujeny, “Filter bank spectrum sensing for cognitive radios,” *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1801–1811, 2008.
- [4] N. Michailow, M. Matth e, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, “Generalized frequency division multiplexing for 5th generation cellular networks,” *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3045–3061, 2014.
- [5] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J.-F. Frigon, “Universal-filtered multi-carrier technique for wireless systems beyond lte,” in *IEEE Globecom Workshop*, 2013, pp. 223–228.
- [6] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [7] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G cellular: It will work!” *IEEE Access*, vol. 1, pp. 335–349.
- [8] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in *IEEE 77th vehicular technology conference (VTC Spring)*, 2013, pp. 1–5.
- [9] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019.
- [10] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” *IEEE Communications Magazine*, 2019.
- [11] E. Basar, M. Di Renzo, J. De Rosny, M. Debbah, M.-S. Alouini, and R. Zhang, “Wireless communications through reconfigurable intelligent surfaces,” *IEEE Access*, vol. 7, pp. 116 753–116 773, 2019.
- [12] T. H. Lee and A. Hajimiri, “Oscillator phase noise: A tutorial,” *IEEE Journal of Solid State Circuits*, vol. 35, no. 3, pp. 326–336, 2000.

- [13] A. Hajimiri and T. H. Lee, “A general theory of phase noise in electrical oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2, pp. 179–194, 1998.
- [14] B. Razavi, “A study of phase noise in CMOS oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 31, no. 3, pp. 331–343, 1996.
- [15] M. R. Khanzadi, D. Kuylenstierna, A. Panahi, T. Eriksson, and H. Zirath, “Calculation of the performance of communication systems from measured oscillator phase noise,” *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 61, no. 5, pp. 1553–1565, 2014.
- [16] T. Pollet, M. Van Bladel, and M. Moeneclaey, “BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise,” *IEEE Transactions on Communications*, vol. 43, no. 234, pp. 191–193, 1995.
- [17] R. Tkach and A. Chraplyvy, “Phase noise and linewidth in an InGaAsP DFB laser,” *IEEE Journal of Lightwave Technology*, vol. 4, no. 11, pp. 1711–1716, 1986.
- [18] L. Tomba, “On the effect of Wiener phase noise in OFDM systems,” *IEEE Transactions on Communications*, vol. 46, no. 5, pp. 580–583, 1998.
- [19] G. Colavolpe, A. Barbieri, and G. Caire, “Algorithms for iterative decoding in the presence of strong phase noise,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, 2005.
- [20] A. Hajimiri, S. Limotyrakis, and T. H. Lee, “Jitter and phase noise in ring oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 34, no. 6, pp. 790–804, 1999.
- [21] A. A. Abidi, “Phase noise and jitter in CMOS ring oscillators,” *IEEE Journal of Solid-State Circuits*, vol. 41, no. 8, pp. 1803–1816, 2006.
- [22] M. Moeneclaey, “The effect of synchronization errors on the performance of orthogonal frequency-division multiplexed (OFDM) systems,” in *Emerging Techniques for Communication Terminals*, 1997.
- [23] J. A. Bingham, “Multicarrier modulation for data transmission: An idea whose time has come,” *IEEE Communications Magazine*, vol. 28, no. 5, pp. 5–14, 1990.
- [24] B. Muquet, Z. Wang, G. B. Giannakis, M. De Courville, and P. Duhamel, “Cyclic prefixing or zero padding for wireless multicarrier transmissions?” *IEEE Transactions on Communications*, vol. 50, no. 12, pp. 2136–2148, 2002.
- [25] L. Deneire, B. Gyselinckx, and M. Engels, “Training sequence versus cyclic prefix—a new look on single carrier communication,” *IEEE Communications Letters*, vol. 5, no. 7, pp. 292–294, 2001.

- [26] O. Rousseaux, G. Leus, and M. Moonen, “Estimation and equalization of doubly selective channels using known symbol padding,” *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 979–990, 2006.
- [27] H. Steendam and M. Moeneclaey, “Different guard interval techniques for OFDM: performance comparison,” in *Multi-Carrier Spread Spectrum 2007*. Springer, 2007, pp. 11–24.
- [28] A. G. Armada, “Understanding the effects of phase noise in orthogonal frequency division multiplexing (OFDM),” *IEEE Transactions on Broadcasting*, vol. 47, no. 2, pp. 153–159, 2001.
- [29] A. G. Armada and M. Calvo, “Phase noise and sub-carrier spacing effects on the performance of an OFDM communication system,” *IEEE Communications Letters*, vol. 2, no. 1, pp. 11–13, 1998.
- [30] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [31] A. P. Worthen and W. E. Stark, “Unified design of iterative receivers using factor graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 843–849, 2001.
- [32] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., 1988.
- [33] J. S. Yedidia, W. T. Freeman, and Y. Weiss, “Constructing free-energy approximations and generalized belief propagation algorithms,” *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.
- [34] M. J. Wainwright, M. I. Jordan *et al.*, “Graphical models, exponential families, and variational inference,” *Foundations and Trends in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.
- [35] T. P. Minka, “Expectation propagation for approximate bayesian inference,” in *Proceedings of the Seventeenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 2001, pp. 362–369.
- [36] E. Riegler, G. E. Korkelund, C. N. Manchón, M.-A. Badiu, and B. H. Fleury, “Merging belief propagation and the mean field approximation: A free energy approach,” *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 588–602, 2013.
- [37] K. P. Murphy, Y. Weiss, and M. I. Jordan, “Loopy belief propagation for approximate inference: An empirical study,” in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1999, pp. 467–475.

- [38] Y. Weiss and W. T. Freeman, “On the optimality of solutions of the max-product belief-propagation algorithm in arbitrary graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 736–744, 2001.
- [39] M. Pretti, “A message-passing algorithm with damping,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, no. 11, 2005.
- [40] T. Heskes *et al.*, “Stable fixed points of loopy belief propagation are minima of the Bethe free energy,” *Advances in neural information processing systems*, vol. 15, pp. 359–366, 2003.
- [41] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [42] Y. Mao and A. H. Banihashemi, “Decoding low-density parity-check codes with probabilistic schedule,” in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, vol. 1, 2001, pp. 119–123.
- [43] E. Yeo, P. Pakzad, B. Nikolic, and V. Anantharam, “High throughput low-density parity-check decoder architectures,” in *IEEE Global Telecommunications Conference*, vol. 5, 2001, pp. 3019–3024.
- [44] J. Zhang and M. P. Fossorier, “Shuffled iterative decoding,” *IEEE Transactions on Communications*, vol. 53, no. 2, pp. 209–213, 2005.
- [45] H. Kfir and I. Kanter, “Parallel versus sequential updating for belief propagation decoding,” *Statistical Mechanics and its Applications*, vol. 330, no. 1, pp. 259–270, 2003.
- [46] M. M. Mansour and N. R. Shanbhag, “High-throughput LDPC decoders,” *IEEE transactions on very large scale integration (VLSI) Systems*, vol. 11, no. 6, pp. 976–996, 2003.
- [47] D. E. Hocevar, “A reduced complexity decoder architecture via layered decoding of LDPC codes,” in *IEEE Workshop on Signal Processing Systems*, 2004, pp. 107–112.
- [48] E. Sharon, S. Litsyn, and J. Goldberger, “An efficient message-passing schedule for LDPC decoding,” in *IEEE Convention of Electrical and Electronics Engineers in Israel*, 2004, pp. 223–226.
- [49] —, “Efficient serial message-passing schedules for LDPC decoding,” *IEEE Transactions on Information Theory*, vol. 53, no. 11, pp. 4076–4091, 2007.
- [50] G. Elidan, I. McGraw, and D. Koller, “Residual belief propagation: Informed scheduling for asynchronous message passing,” *arXiv preprint arXiv:1206.6837*, 2012.

- [51] A. I. V. Casado, M. Griot, and R. D. Wesel, “Informed dynamic scheduling for belief-propagation decoding of LDPC codes,” in *IEEE International Conference on Communications*, 2007, pp. 932–937.
- [52] C. A. Aslam, Y. L. Guan, K. Cai, and G. Han, “Low-complexity belief-propagation decoding via dynamic silent-variable-node-free scheduling,” *IEEE Communications Letters*, vol. 21, no. 1, pp. 28–31, 2017.
- [53] Y. Weiss and W. T. Freeman, “Correctness of belief propagation in Gaussian graphical models of arbitrary topology,” in *Advances in neural information processing systems*, 2000, pp. 673–679.
- [54] D. J. MacKay and R. M. Neal, “Near Shannon limit performance of low density parity check codes,” *IET Electronics letters*, vol. 32, no. 18, pp. 1645–1646, 1996.
- [55] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, “Turbo decoding as an instance of Pearl’s “belief propagation” algorithm,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 140–152, 1998.
- [56] M.-A. Badiu, G. E. Kirlund, C. N. Manchón, E. Riegler, and B. H. Fleury, “Message-passing algorithms for channel estimation and decoding using approximate inference,” in *International Symposium on Information Theory (ISIT)*, 2012, pp. 2376–2380.
- [57] M.-A. Badiu, C. N. Manchón, and B. H. Fleury, “Message-passing receiver architecture with reduced-complexity channel estimation,” *IEEE Communications Letters*, vol. 17, no. 7, pp. 1404–1407, 2013.
- [58] J. Dauwels and H.-A. Loeliger, “Phase estimation by message passing,” in *IEEE International Conference on Communications*, vol. 1, 2004, pp. 523–527.
- [59] A. Barbieri, G. Colavolpe, and G. Caire, “Joint iterative detection and decoding in the presence of phase noise and frequency offset,” *IEEE Transactions on Communications*, vol. 55, no. 1, pp. 171–179, 2007.
- [60] F. Z. Merli and G. M. Vitetta, “A factor graph approach to the iterative detection of OFDM signals in the presence of carrier frequency offset and phase noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 868–877, 2008.
- [61] H. Wymeersch, J. Lien, and M. Z. Win, “Cooperative localization in wireless networks,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, 2009.
- [62] S. Zarrin and T. J. Lim, “Belief propagation on factor graphs for cooperative spectrum sensing in cognitive radio,” in *IEEE Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2008, pp. 1–9.

- [63] H. Wymeersch, F. Penna, and V. Savić, “Uniformly reweighted belief propagation for estimation and detection in wireless networks,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1587–1595, 2012.
- [64] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, 2009.
- [65] D. Baron, S. Sarvotham, and R. G. Baraniuk, “Bayesian compressive sensing via belief propagation,” *IEEE Transactions on Signal Processing*, vol. 58, no. 1, pp. 269–280, 2010.
- [66] S. Rangan, “Generalized approximate message passing for estimation with random linear mixing,” in *IEEE International Symposium on Information Theory (ISIT)*, 2011, pp. 2168–2172.
- [67] H. Wymeersch, F. Penna, and V. Savić, “Uniformly reweighted belief propagation: A factor graph approach,” in *IEEE International Symposium on Information Theory (ISIT)*, 2011, pp. 2000–2004.
- [68] L. Barletta, M. Magarini, and A. Spalvieri, “The information rate transferred through the discrete-time Wiener’s phase noise channel,” *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1480–1486, 2012.
- [69] ———, “Estimate of information rates of discrete-time first-order Markov phase noise channels,” *IEEE Photonics Technology Letters*, vol. 23, no. 21, pp. 1582–1584, 2011.
- [70] S. H. Rezenom and F. Takawira, “Iterative low complexity algorithm for LDPC systems in the presence of phase noise,” *IEEE Wireless Communication Letters*, vol. 6, no. 6, pp. 794–797, 2017.
- [71] L. Barletta, “Computing the information rate of discrete-time Wiener phase noise channels by parametric bayesian tracking,” *Optics Express*, vol. 24, no. 23, pp. 26 612–26 617, 2016.

II

Papers

Paper 1

Fast Detection of OFDM Systems Using Graphical Models

in *Proceedings of IEEE Signal Processing Workshop*, pp. 154–158, UT, USA, 2015.

Abstract

In this paper, we investigate the effect of four different message schedules on the performance of an OFDM receiver with unknown carrier frequency and phase noise offsets. One of the methods uses a serial schedule and the remaining three techniques use non-serial message schedules. The serial schedule is shown to converge in only one iteration. The results also show that fast graphical estimators can be designed by using non-serial message schedules with damping to approach the bit error rates of the serial schedule while reducing the computational time for convergence. In particular, the damped flooding message schedule using four iterations reduces the computational time for convergence by more than 30 % compared to the serial message schedule using one iteration for signal to noise ratios lower than 20 dB.

I Introduction

Graphical models are becoming increasingly attractive as a research tool for inference problems in communications receivers. The work by Gallager [1] to represent code bits using graphs in the development of the Low Density Parity Check codes (LDPC) is the first to introduce graphical models in communications systems. The use of graphical models has since then showed significant improvement especially within the context of decoding capacity approaching codes. The work of Forney [2], Tanner [3] and Wiberg [4] contributed towards understanding the modelling of communication receivers using graphical models. Kschischang in [5] presents a good review of graphical models in their various manifestations - Markov random fields, factor graphs and Bayesian networks.

Graphical models have also gained research interest in other areas of communication

receivers - joint channel estimation and decoding problems [6]–[9], decoding of LDPC systems in the presence of phase noise [10], [11], joint carrier frequency offset and phase noise estimation in Orthogonal Frequency Division Multiplexing (OFDM) systems [12], cooperative wireless positioning problems [13] - to mention just a few.

Graphical models arise out of the factorisation of the probability density function of the underlying problem. The estimates can theoretically be found by using maximum likelihood detection, which may not be practical for complex problems. In graphical models, the inferred parameters are obtained in a distributed manner. The significance of graphical models in communications problems is therefore largely due to the reduction in computational complexity. To obtain the inferred parameters, message passing algorithms are derived that minimise the Kullback-Liebler divergence of the underlying density function of the problem. One such message passing method is the belief propagation [14] and its widely used implementation - the sum-product algorithm [15]. The link between the belief propagation technique and the decoding of LDPC and turbo codes is demonstrated in [16], [17].

Once the underlying graphical model has been developed, messages are constructed that are then propagated in the graph using the belief propagation algorithm [15]. However, since most graphical models contain cycles, the use of a message passing schedule is important. The first work to question whether message passing schedules affect performance was Forney [18]. Several authors have developed different message passing schedules such as the two-way schedule [5], the flooding schedule [5], probabilistic schedule [19], shuffled belief propagation schedule [20], sequential schedule [21], [22]. The work in [22] showed in particular that the sequential schedule converges in half the total number of iterations compared to the flooding schedule. The main difference between these scheduling methods is the order of message update between the variable and factor nodes of the graph. All of these schedules have been developed and tested within the context of LDPC decoding.

However, unlike decoding capacity approaching codes whose graphical models are characterised with minimal node density and discrete density functions, the effect of scheduling on communication receivers with dense nodes and continuous probability density functions has not been extensively studied.

In this paper, we show that the use of a simple schedule using belief propagation algorithm converges in only one iteration when applied to the estimation problem of carrier frequency offset and phase noise in an OFDM system [12]. We also show that other message passing schedules can be constructed to optimise the convergence time. Optimising the performance of the BP algorithm graphically in the manner described in this paper is shown to provide some performance advantages.

The paper is organised as follows. Section II discusses the OFDM system and the resulting graphical model. Section III presents the message construction techniques. Section IV presents the simulation parameters and presents the results. Finally, conclusions are drawn in Section V.

Notation: We use $\mathcal{CN}(0, \sigma^2)$ to denote a complex Gaussian distribution with zero

mean and variance σ^2 . For matrix operations, $\mathbb{C}^{N \times M}$ denotes a matrix of size $N \times M$ with complex entries and $(\cdot)^H$ denotes the Hermitian transpose of a given matrix.

II System Model

A OFDM System Model

We consider an OFDM system with N subcarriers. The user signal $\mathbf{d} = \{d_0, d_1, \dots, d_{N-1}\}$ is converted to parallel and applied to the Inverse Discrete Fourier Transform (IDFT). After adding a cyclic prefix of N_{CP} and appropriate filtering operations, the signal is transmitted. The received signal at the receiver, after removal of the cyclic prefix, can be expressed as

$$\mathbf{y} = \Gamma(v, \phi) \mathbf{F}^H \mathbf{D} \mathbf{F}_L \mathbf{h} + \mathbf{w}, \quad (1)$$

where v is the carrier frequency offset and the phase noise is

$\phi = [\phi_0, \phi_1, \dots, \phi_{N-1}]$. The phase noise process is distributed according to a Gaussian Markov process, i.e.,

$$\phi_n = \phi_{n-1} + \Delta\phi, \quad (2)$$

with ϕ_n denoting the phase noise at the n^{th} subcarrier of the OFDM symbol and $\Delta\phi \sim \mathcal{N}(0, \sigma_\phi^2)$, where σ_ϕ^2 is the variance of the phase noise, and is computed as $\sigma_\phi^2 = \frac{\gamma}{N}$ where γ is the Wiener process variance [23]. Furthermore $\Gamma(v, \phi) = \text{diag}(e^{j\phi_0}, e^{j2\pi v/N + j\phi_1}, \dots, e^{j2\pi v(N-1)/N + j\phi_{N-1}})$, $\mathbf{D} = \text{diag}(d_0, d_1, \dots, d_{N-1})$, $\mathbf{F} \in \mathbb{C}^{N \times N}$ is the DFT operator with entries $\frac{1}{\sqrt{N}} \exp(-j2\pi pq/N)$ and $p, q \in [0, N-1]$, and $\mathbf{F}_L \in \mathbb{C}^{N \times L}$ is the matrix constructed by taking the first L columns of matrix \mathbf{F} . The vector \mathbf{w} represents the Additive White Gaussian Noise (AWGN) distributed according to $\mathcal{CN}(0, \sigma^2)$ where σ^2 is the variance of the noise. The channel impulse response is $\mathbf{h} = [h_0, h_1, \dots, h_{L-1}]$ with h_l representing the channel coefficients at the l^{th} path which are distributed according to a wide-sense stationary complex Gaussian process, i.e., $h_l \sim \mathcal{CN}(0, \sigma_h^2(l))$, where $\sigma_h^2(l)$ is the variance of path l and L is the channel length.

B Graphical Model

The main problem is finding estimates for the *a posteriori* probability density function

$$\left(\hat{v}, \hat{\phi} \right) = \underset{v, \phi}{\text{argmax}} \quad p(v, \phi | \mathbf{y}, \mathbf{d}, \mathbf{h}). \quad (3)$$

Therefore

$$p(v, \phi | \mathbf{y}, \mathbf{d}, \mathbf{h}) \propto p(\mathbf{y} | v, \phi, \mathbf{d}, \mathbf{h}) p(v) p(\phi), \quad (4)$$

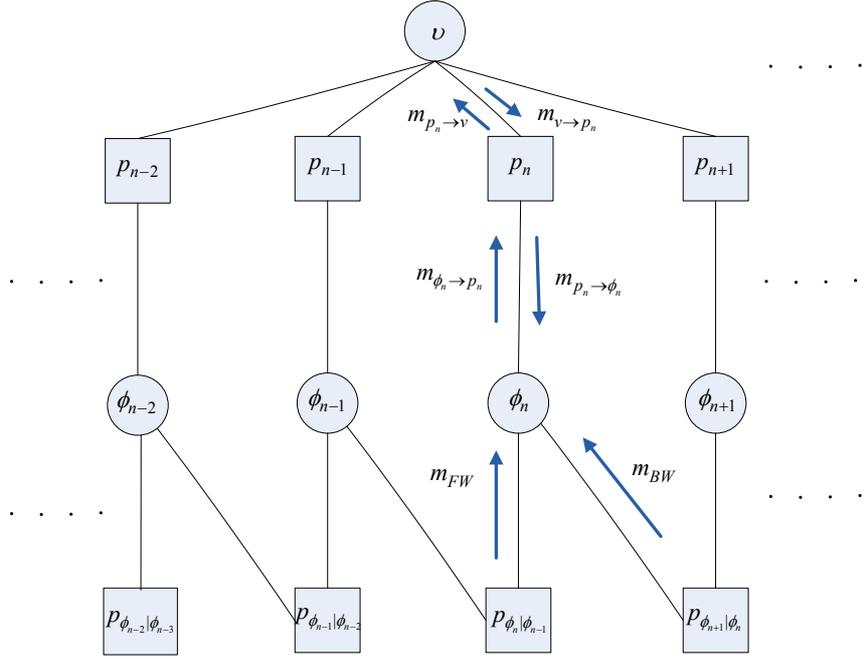


Figure 1: Graphical model.

The data and channel frequency response are assumed to be known at the receiver. Therefore, $\mathbf{t} = \mathbf{F}^H \mathbf{D} \mathbf{F}_L \mathbf{h}$ is known at the receiver. The density functions can be factorised as [12]

$$p(v, \phi | \mathbf{y}, \mathbf{t}) \propto p(v) p(\phi_0) \prod_{n=1}^{N-1} p(\phi_n | \phi_{n-1}) \prod_{n=0}^{N-1} p(y_n | v, \phi_n, t_n), \quad (5)$$

where $p(y_n | v, \phi_n, t_n) = \mathcal{CN}(e^{j2\pi v n/N + j\phi_n t_n}, \sigma^2)$ and $p(\phi_n | \phi_{n-1}) = \mathcal{N}(0, \sigma_\phi^2)$. A graphical model can then be constructed [12] for the factorised joint probability density function given in (5) and is shown in Figure 1.

III Message Construction

For a given graph, different message passing schedules can be defined [15]. The most commonly used is the flooding technique [5], [15] whereby all variable and factor nodes update their messages at the same time. Another message passing schedule where the nodes sequentially update their messages is described in [22]. Other message schedules can be defined that are a combination of these two schedules. In this work, we define four different message passing schedules and are shown in Table 1 in detail. The flooding schedule, denoted by ‘Schedule f’, updates the messages from the factor and variable nodes simultaneously. The scheme denoted by ‘Schedule s’ progressively updates the messages in one direction, and when these updates are complete, the messages in the reverse direction are updated. We will refer to this schedule as serial schedule owing the simultaneous update of messages in one direction. The scheme denoted by ‘Schedule c’ computes three sets of

messages in parallel in each iteration before formulating a belief. Similarly, the message scheme denoted by ‘Schedule x’ alternates between the simultaneous update of messages from the factor to variables nodes followed by the subsequent simultaneous update of the messages from the variable nodes to the factor nodes. Therefore ‘Schedule c’ and ‘Schedule x’ can be looked at as some combination of the serial and parallel schedules. We refer to the ‘Schedule c’, ‘Schedule x’ and ‘Schedule f’ as non-serial schedules.

In this work, we use message damping, where a message is updated according to a linear combination of the messages from the current and previous iterations [24]. A damping parameter $\alpha \in (0, 1]$ is used and the messages are updated according to $m^{(t)}(x_i) = \alpha m^{(t)}(x_i) + (1 - \alpha)m^{(t-1)}(x_i)$ where t indicates the iteration index and the message $m(x_i)$ is in the logarithmic domain. All the messages shown in Figure 1 are computed using the belief propagation algorithm [12], [15].

IV Results

The performance of the system is investigated through simulations. The number of subcarriers used is 256, with 31 of these set to be null subcarriers. The length of cyclic prefix is 30. The transmitted data is modulated using QPSK. Virtual subcarriers are used at both ends of the transmitted data in order to reduce inter-carrier interference. The channel is assumed static for the duration of the OFDM symbol and changes according to the variance $\sigma^2(l) = \frac{1-e^{-1/5}}{1-e^{-L/5}}e^{-l/5}$ with $l = 0, 1, \dots, 15$. The frequency offset v is selected from a uniform distribution in the range $[-0.2, 0.2]$ and is fixed for the duration of the OFDM symbol. The phase noise is constructed as follows. The first phase noise sample ϕ_0 of every symbol is selected from a uniform distribution in the range $[-\pi/2, \pi/2]$, and the subsequent samples are generated according to (2). The Wiener process variance γ is assumed to be 0.1. The bit-error rate (BER) is used as a performance metric at various signal-to-noise ratios (SNR) shown in dBs. Perfect time synchronisation is considered at the receiver. Therefore, after sampling, the received sequences are applied to the iterative receiver, which has knowledge of the transmitted data and channel parameters. The receiver creates a sampling space of the unknown parameters. The sampling space of the frequency offset Q_v is in $[-0.5, 0.5]$ and that of the phase noise Q_ϕ is in $[-\pi, \pi]$. In both cases, 25 samples are used to represent the carrier frequency offset and phase noise processes.

The max-sum implementation of the belief propagation is used in the logarithmic domain. Damping has been found to improve the bit error rates at high SNRs. A factor of 0.2 has been used as damping parameter for the non-serial message schedules because it gives the lowest error rates. The serial schedule shows no improvement as a result of using damped messages. Figure 2 compares the error rate performance of the different schedules at SNR of 30 dB. It is interesting to note that the serial schedule converges within the first iteration. In comparison, the error rate of ‘Schedule c’ stabilises after the fourth iteration, while that of the ‘Schedule f’ and ‘Schedule x’ converges at the sixth iteration. Figure 3

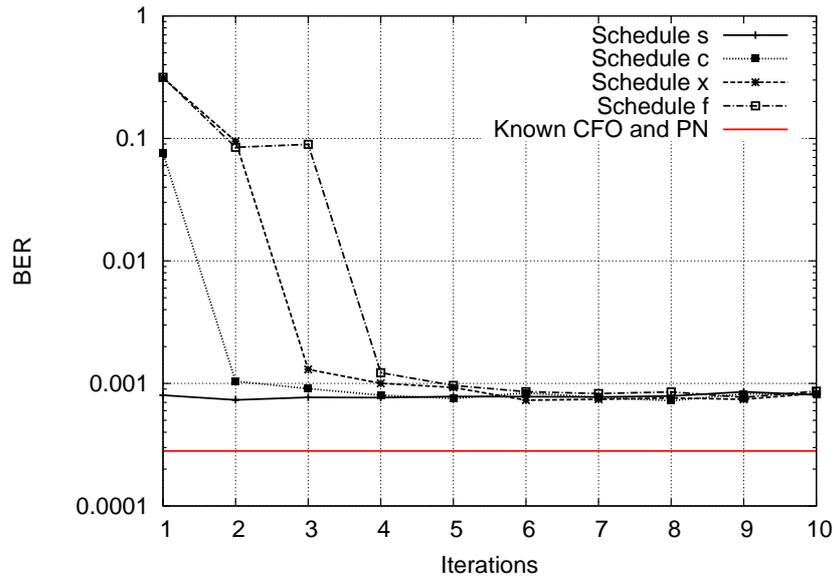


Figure 2: Convergence performance of different schedules at 30 dB.

shows the BER performance for various SNR values for selected number of iterations. The serial schedule performs consistently for all SNR levels. This difference is due to the use of recently updated messages as an argument to the computation of the subsequent messages within the same iterative cycle. The serial schedule, by its sequential nature, uses the most updated information in each iterative cycle compared to other schedules. To illustrate further, for each iteration as shown in Table 1, five of the messages in ‘Schedule s’ use recently updated messages. In contrast, four of the messages in ‘Schedule c’ use recent updates and only two are used in ‘Schedule x’. None of the messages in the flooding schedule use recent message updates in the same iterative cycle. The updated messages are only used in the subsequent iterative cycle.

The differences in error rates has practical implications. The results show that the flooding schedule using four iterations achieves similar error rates as the serial schedule for SNR values less than 20 dB. This dynamics can be optimised to reduce the computational time as the flooding schedule using four iterations only needs four computational time units to achieve the same error rates as the serial one, which requires six computational time units. Under this assumption of parallelism, the flooding schedule reduces the computational time by more than 30 % compared with the serial schedule that uses one iteration. Similarly, for SNR values larger than 20 dB, the ‘Schedule c’ using five iterations achieves similar error rates as the serial schedule. However, its computational time would be more than doubled in comparison to the serial schedule. Therefore, the serial schedule gives the best performance in terms of the computational time and error rates for SNR values larger than 20 dB.

Table 1: Message Construction. $F\{x\}$ represents functional dependence on the message x . The message sequencing within each iteration is denoted with 'P'. For example, P1 indicates the time slot for executing a particular message. Two messages with P1 indicate that the messages are computed in parallel.

Schedule c	Schedule f
Initialisation $m_{p_n \rightarrow v}^{(t-1)} = 1$ $m_{p_n \rightarrow \phi_n}^{(t-1)}(\phi_n) = 1$ $m_{FW}^{(t-1)}(\phi_n) = 1$ $m_{BW}^{(t-1)}(\phi_n) = 1$ P1: $m_{\phi_n \rightarrow p_n}^{(t)} = F\{m_{FW}^{(t-1)}, m_{BW}^{(t-1)}\}$ P1: $m_{v \rightarrow p_n}^{(t)}(v) = F\{m_{p_n \rightarrow v}^{(t-1)}\}$ P2: $m_{p_n \rightarrow \phi_n}^{(t)}(\phi_n) = F\{m_{v \rightarrow p_n}^{(t)}\}$ P2: $m_{p_n \rightarrow v}^{(t)}(v) = F\{m_{\phi_n \rightarrow p_n}^{(t)}\}$ P3: $m_{FW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t)}\}$ P3: $m_{BW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t)}\}$	Initialisation $m_x^{(t-1)} = 1, \text{ for all messages } x$ P1: $m_{p_n \rightarrow \phi_n}^{(t)}(\phi_n) = F\{m_{v \rightarrow p_n}^{(t-1)}\}$ P1: $m_{p_n \rightarrow v}^{(t)}(v) = F\{m_{\phi_n \rightarrow p_n}^{(t-1)}\}$ P1: $m_{FW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t-1)}\}$ P1: $m_{BW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t-1)}\}$ P1: $m_{\phi_n \rightarrow p_n}^{(t)} = F\{m_{FW}^{(t-1)}, m_{BW}^{(t-1)}\}$ P1: $m_{v \rightarrow p_n}^{(t)}(v) = F\{m_{p_n \rightarrow v}^{(t-1)}\}$
Schedule x	Schedule s
Initialisation $m_{p_n \rightarrow v}^{(t-1)}(v) = 1$ $m_{p_n \rightarrow \phi_n}^{(t-1)}(\phi_n) = 1$ $m_{FW}^{(t-1)}(\phi_n) = 1$ $m_{BW}^{(t-1)}(\phi_n) = 1$ P1: $m_{\phi_n \rightarrow p_n}^{(t)} = F\{m_{FW}^{(t-1)}, m_{BW}^{(t-1)}\}$ P1: $m_{v \rightarrow p_n}^{(t)}(v) = F\{m_{p_n \rightarrow v}^{(t-1)}\}$ P2: $m_{FW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t-1)}\}$ P2: $m_{BW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t-1)}\}$ P2: $m_{p_n \rightarrow v}^{(t)}(v) = F\{m_{\phi_n \rightarrow p_n}^{(t)}\}$ P2: $m_{p_n \rightarrow \phi_n}^{(t)}(\phi_n) = F\{m_{v \rightarrow p_n}^{(t)}\}$	Initialisation $m_{v \rightarrow p_n}^{(t-1)}(v) = 1$ P1: $m_{p_n \rightarrow \phi_n}^{(t)}(\phi_n) = F\{m_{v \rightarrow p_n}^{(t-1)}\}$ P2: $m_{FW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t)}\}$ P3: $m_{BW}^{(t)}(\phi_n) = F\{m_{p_n \rightarrow \phi_n}^{(t)}\}$ P4: $m_{\phi_n \rightarrow p_n}^{(t)} = F\{m_{FW}^{(t)}, m_{BW}^{(t)}\}$ P5: $m_{p_n \rightarrow v}^{(t)}(v) = F\{m_{\phi_n \rightarrow p_n}^{(t)}\}$ P6: $m_{v \rightarrow p_n}^{(t)}(v) = F\{m_{p_n \rightarrow v}^{(t)}\}$

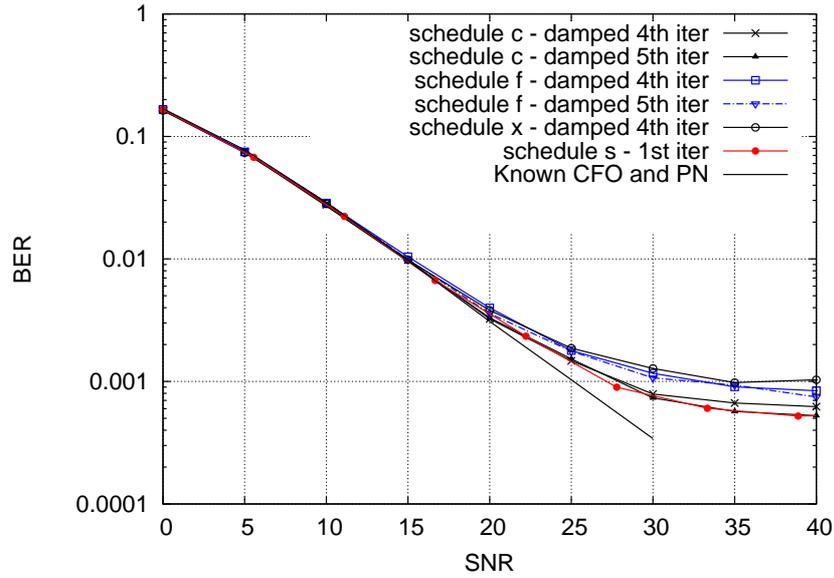


Figure 3: Error rate performance versus SNR for the different schedules.

V Conclusion

In this work, we have considered the problem of joint estimation of carrier frequency offset and phase noise in an OFDM system. We have used four different schedules using belief propagation and found that the serial schedule converges in one iteration. The flooding schedule is shown to converge faster than the serial schedule for signal to noise ratios lower than 20 dB. This dynamics of message scheduling may be exploited to design efficient and fast estimators for problems which can be represented using graphical models.

References

- [1] R. G. Gallager, “Low-density parity-check codes,” *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, 1962.
- [2] G. Forney Jr, “The forward-backward algorithm,” in *Proceedings of the 34th Allerton Conference on Communications, Control and Computing*, 1996, pp. 432–446.
- [3] R. M. Tanner, “A recursive approach to low complexity codes,” *IEEE Transactions on Information Theory*, vol. 27, no. 5, pp. 533–547, 1981.
- [4] N. Wiberg, “Codes and decoding on general graphs,” Ph.D. dissertation, Linköping University, 1996.
- [5] F. R. Kschischang and B. J. Frey, “Iterative decoding of compound codes by probability propagation in graphical models,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 219–230, 1998.
- [6] Y. Liu, L. Brunel, and J. J. Boutros, “Joint channel estimation and decoding using Gaussian approximation in a factor graph over multipath channel,” in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, 2009, pp. 3164–3168.
- [7] M.-A. Badiu, G. E. Kirkelund, C. N. Manchón, E. Riegler, and B. H. Fleury, “Message-passing algorithms for channel estimation and decoding using approximate inference,” in *IEEE International Symposium on Information Theory (ISIT)*, 2012, pp. 2376–2380.
- [8] C. Knievel, P. A. Hoeher, A. Tyrrell, and G. Auer, “Multi-dimensional graph-based soft iterative receiver for MIMO-OFDM,” *IEEE Transactions on Communications*, vol. 60, no. 6, pp. 1599–1609, 2012.
- [9] P. Schniter, “A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 8, pp. 1462–1474, 2011.
- [10] J. Dauwels and H.-A. Loeliger, “Phase estimation by message passing,” in *IEEE International Conference on Communications*, vol. 1, 2004, pp. 523–527.
- [11] G. Colavolpe, A. Barbieri, and G. Caire, “Algorithms for iterative decoding in the presence of strong phase noise,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, 2005.
- [12] F. Z. Merli and G. M. Vitetta, “A factor graph approach to the iterative detection of OFDM signals in the presence of carrier frequency offset and phase noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 868–877, 2008.

- [13] H. Wymeersch, J. Lien, and M. Z. Win, “Cooperative localization in wireless networks,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, 2009.
- [14] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [15] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [16] D. J. MacKay and R. M. Neal, “Near Shannon limit performance of low density parity check codes,” *IET Electronics letters*, vol. 32, no. 18, pp. 1645–1646, 1996.
- [17] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, “Turbo decoding as an instance of pearl’s “belief propagation” algorithm,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 140–152, 1998.
- [18] G. D. Forney Jr, “On iterative decoding and the two-way algorithm,” in *International Symposium on Turbo Codes and Related Topics*, 1997, pp. 12–15.
- [19] Y. Mao and A. H. Banihashemi, “Decoding low-density parity-check codes with probabilistic schedule,” in *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, vol. 1, 2001, pp. 119–123.
- [20] J. Zhang and M. Fossorier, “Shuffled belief propagation decoding,” in *IEEE Asilomar Conference on Signals, Systems and Computers*, vol. 1, 2002, pp. 8–15.
- [21] H. Kfir and I. Kanter, “Parallel versus sequential updating for belief propagation decoding,” *Statistical Mechanics and its Applications*, vol. 330, no. 1, pp. 259–270, 2003.
- [22] E. Sharon, N. Presman, and S. Litsyn, “Analysis of LDPC decoding schedules,” in *IEEE International Symposium on Information Theory (ISIT)*, 2009, pp. 1669–1673.
- [23] T. Pollet, M. van Bladel, and M. Moeneclaey, “BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise,” *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 191–193, Feb./Mar./Apr. 1998.
- [24] M. Pretti, “A message-passing algorithm with damping,” *Journal of Statistical Mechanics: Theory and Experiment*, no. 11, 2005.

Paper 2

Iterative Low Complexity Algorithm for LDPC Systems in the Presence of Phase Noise

IEEE Wireless Communications Letters, no. 6, pp. 794–797, 2017.

Abstract

In this paper, we propose a low-complexity graph based iterative receiver for LDPC coded systems in the presence of strong phase noise. The proposed receiver exploits the inherent circular characteristics of phase noise. The graphical messages are then constructed using complex Gaussian density functions. We show the proposed method achieves similar bit error rate performance as the existing lowest complexity algorithm, and achieves lower complexity for higher modulation.

I Introduction

The error performance of recent communication systems have improved dramatically. This improvement has been made possible largely by using capacity approaching codes such as the low density parity check codes (LDPC) [1]. However, phase noise has been known to severely affect the performance of communication systems. Therefore, effective algorithms have to be designed to tackle the effect of phase noise in coded communication systems. Several authors have proposed algorithms to detect coded data in the presence of phase noise [2]–[4].

The iterative detection of an LDPC communication system in a phase noise channel has been presented in [2] where the authors derive a graphical model from the posterior probability density of the system. The messages in the graph are then constructed and propagated using the sum product algorithm [5]. The use of graphical models in communication systems has been described in detail in [5],[6]. In most cases, for continuous random

variables, there is no explicit analytical solution to the messages constructed by the sum product algorithm and hence many authors resort to approximations. One such approximation is discussed in [2] where the continuous phase noise messages are approximated by collecting probabilities computed at discrete samples in the phase domain. The error performance of the discrete message based methods generally improve with the number of samples. There is a threshold in the number of samples used, beyond which the error rate improves minimally [7].

The computational complexity of the discrete method increases with the number of discrete samples. In order to reduce this complexity, the authors in [2] approximated the continuous phase noise messages using the Tikhonov probability density function. Hence, instead of the discrete messages, the parameters of the Tikhonov probability density function are propagated in the graph. The authors show this approach reduces the computational complexity significantly while achieving comparable performance to the discrete messages.

In a similar work, the authors in [8] approximate the phase messages in the graph using multiple Tikhonov probability density functions and then reduce the number of functions using mixture reduction techniques. They show this approach improves the error performance compared to the method in [2] under the parameters used in their work. This approach has slightly more computational complexity than the method in [2] but significantly lower than the discrete messages.

Existing methods have so far modelled the phase noise random variables in the linear domain, and have not exploited the fact that phase noise is a complex random variable defined over the unit circle. The modelling of phase noise as a circular random variable requires a new framework in constructing and propagating messages in the graphical model. In this work, we study the iterative detection of an LDPC communication system in phase noise channels [2]. We deviate from the existing approach of restricting the phase noise space to the linear domain. Instead, we transform the phase noise space to the complex domain and construct all messages in the graphical model. We show that this method achieves error performance that is comparable to the existing low-complexity algorithm, while reducing the complexity even further.

This paper is structured as follows. Section II presents the graphical model for coded communication systems in the presence of phase noise. Section III presents the proposed method and the framework used to construct messages in the graphical model. Section IV presents the simulation results and finally conclusion is presented in Section V.

II Problem Formulation

We consider a communication system where information bits b_n are coded and subsequently mapped using symbols from a set \mathcal{A} of size M . The coded and mapped symbols

x_n are transmitted through a Wiener phase noise channel. The received signal is

$$y_n = x_n e^{j\phi_n} + w_n, \quad n = 0, \dots, N-1, \quad (1)$$

where the phase process is $\phi_n = \phi_{n-1} + \Delta_{\phi_n}$ with the phase increment Δ_{ϕ_n} selected from the Gaussian density function with zero mean and variance σ_ϕ^2 , with $\phi_0 \sim \mathcal{U}[0, 2\pi)$. The complex additive white Gaussian noise process w_n has zero mean and variance $E[|w_n|^2] = \sigma^2$.

The posterior distribution can be written as

$$p(\mathbf{b}, \mathbf{x}, \phi | \mathbf{y}) \propto p(\mathbf{b}) p(\mathbf{x} | \mathbf{b}) p(\phi) p(\mathbf{y} | \mathbf{x}, \phi) \quad (2)$$

$$\begin{aligned} &\propto p(\mathbf{b}) \mathcal{I}\{\mathbf{x} = \psi(\mathbf{b})\} p(\phi_0) \quad (3) \\ &\times \prod_{n=1}^{N-1} p(\phi_n | \phi_{n-1}) \prod_{n=0}^{N-1} p(y_n | x_n, \phi_n), \end{aligned}$$

where $\psi(\cdot)$ is the encoding and mapping function. The indicator function $\mathcal{I}(\cdot)$ returns a 1 if a valid codeword has been found. Further, $p(y_n | x_n, \phi_n) = \frac{1}{\sqrt{\pi\sigma^2}} \exp\left\{-\frac{|y_n - x_n e^{j\phi_n}|^2}{\sigma^2}\right\}$ and $p(\phi_n | \phi_{n-1}) = \frac{1}{\sqrt{2\pi\sigma_\phi^2}} \exp\left\{-\frac{|\phi_n - \phi_{n-1}|^2}{2\sigma_\phi^2}\right\}$.

A graphical model for the joint probability density function described in (2) has been derived in [2] and is shown in Figure 1. The resulting graph is shown to be composed of the code graph and the phase graph with messages that propagate between these graphs. The message passing procedure has been described in detail in [2]. Here, we briefly outline the construction of the messages shown in the graph using the sum product algorithm [5] as follows. The phase graph first computes the message

$$m_{p_n \rightarrow \phi_n}(\phi_n) = \sum_{x_n \in \mathcal{A}} m_{x_n \rightarrow p_n}(x_n) p(y_n | x_n, \phi_n). \quad (4)$$

The forward message is computed as

$$\begin{aligned} m_{\text{FW}}(\phi_n) &= \int m_{p_{n-1} \rightarrow \phi_{n-1}}(\phi_{n-1}) m_{\text{FW}}(\phi_{n-1}) \quad (5) \\ &\times p(\phi_n | \phi_{n-1}) d\phi_{n-1}, \end{aligned}$$

while the backward message is computed as

$$\begin{aligned} m_{\text{BW}}(\phi_n) &= \int m_{p_{n+1} \rightarrow \phi_{n+1}}(\phi_{n+1}) m_{\text{BW}}(\phi_{n+1}) \quad (6) \\ &\times p(\phi_{n+1} | \phi_n) d\phi_{n+1}. \end{aligned}$$

Finally, the message to the code graph is computed as

$$m_{p_n \rightarrow x_n}(x_n) = \int m_{\text{FW}}(\phi_n) m_{\text{BW}}(\phi_n) p(y_n | x_n, \phi_n) d\phi_n. \quad (7)$$

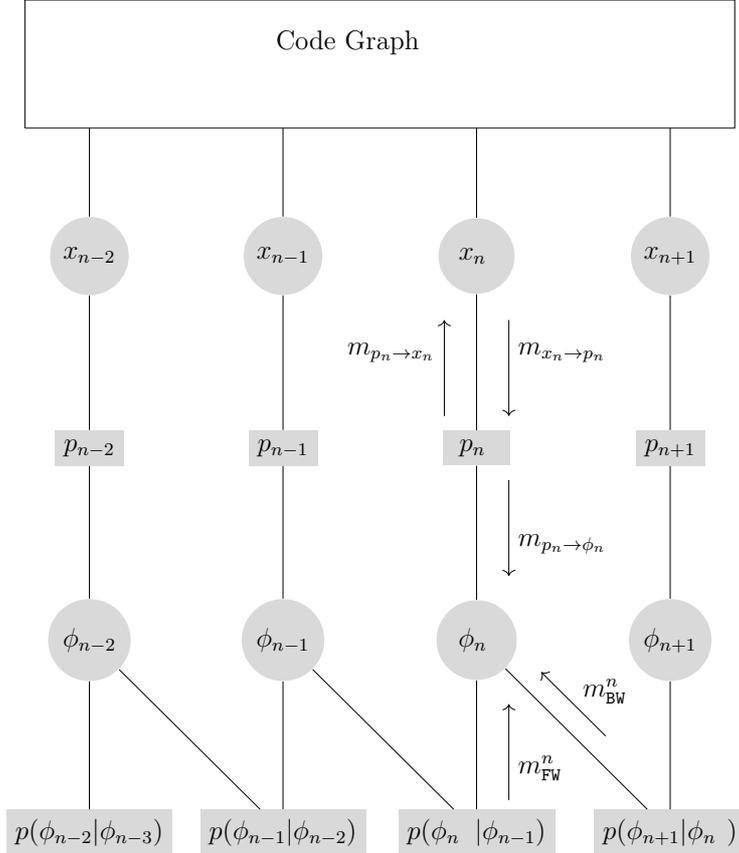


Figure 1: The graphical model showing the code graph and the phase noise graph.

III Proposed Method

We describe the proposed method as follows. The phase noise process in the linear domain is transformed to the complex domain as $X_{\phi_n} = e^{j\phi_n}$. Similarly, the modulated symbols are transformed to the complex domain as $X_n = e^{j\angle x_n}$, where $\angle x_n$ denotes the angle of the modulated symbol. This transformation is a bijection as it preserves the properties of the random variables in the linear domain. The resulting graphical model is similar to Figure 1 with the exception that the phase noise variable nodes ϕ_n are replaced by the circular random variables $X_{\phi_n} = e^{j\phi_n}$. Further, the factor nodes $p(\phi_n|\phi_{n-1})$ are replaced by $p(X_{\phi_n}|X_{\phi_{n-1}})$. Since the phase noise random variables have unity magnitude, we refer this transformation as *circular random variable* method. The messages are constructed as follows using the sum product algorithm.

A Message $m_{X_n \rightarrow p_n}(X_n)$

The first step in the proposed method is to rewrite the message received from the code graph $m_{x_n \rightarrow p_n}(x_n)$ using the framework of circular random variables. We convert the linear domain message $m_{x_n \rightarrow p_n}(x_n)$ to the complex Gaussian message $m_{X_n \rightarrow p_n}(X_n) = \mathcal{CN}(e^{j\angle x_n}; \mathbf{M}_{X_n \rightarrow p_n}, \mathbf{V}_{X_n \rightarrow p_n})$, where $\mathbf{M}_{X_n \rightarrow p_n}$ is the mean of the message $m_{X_n \rightarrow p_n}(X_n)$ and

$\mathbf{V}_{X_n \rightarrow p_n}$ is its variance. This allows the existing discrete message of symbols x_n to be represented as a continuous complex circular distribution.

B Message $m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})$

The message $m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})$ is constructed as

$$m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n}) = \int m_{X_n \rightarrow p_n}(X_n) p(y_n | X_n, X_{\phi_n}) dX_n, \quad (8)$$

where $p(y_n | X_n, X_{\phi_n}) = \mathcal{CN}(y_n; x_n e^{j\phi_n}, \sigma^2)$ is obtained from $p(y_n | x_n, \phi_n)$ using the complex random variables. Therefore, we can write

$$m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n}) = \int \mathcal{CN}(e^{j\angle x_n}; \mathbf{M}_{X_n \rightarrow p_n}, \mathbf{V}_{X_n \rightarrow p_n}) \times \mathcal{CN}\left(e^{j\angle x_n}; \frac{y_n}{e^{j\phi_n}}, \sigma^2\right) de^{j\angle x_n}, \quad (9)$$

where $\mathcal{CN}(y_n; x_n e^{j\phi_n}, \sigma^2) = \mathcal{CN}(e^{j\angle x_n}; \frac{y_n}{e^{j\phi_n}}, \sigma^2)$ has been obtained using properties (A.1) and (A.2) of the Gaussian probability density function presented in the Appendix. Further, using property (A.4) we evaluate (9) as

$$m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n}) \propto \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{p_n \rightarrow X_{\phi_n}}, \mathbf{V}_{p_n \rightarrow X_{\phi_n}}\right), \quad (10)$$

where the mean of $m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})$ is $\mathbf{M}_{p_n \rightarrow X_{\phi_n}} = \left(\frac{\mathbf{M}_{X_n \rightarrow p_n}}{y_n}\right)^*$ and the variance is $\mathbf{V}_{p_n \rightarrow X_{\phi_n}} = \frac{\sigma^2 + \mathbf{V}_{X_n \rightarrow p_n}}{|y_n|^2}$. The conjugate shown in the mean is a result of evaluating the expression $\mathcal{CN}(x^*; \mu, \sigma^2)$.

C Message $m_{\text{FW}}(e^{j\phi_n})$

The forward message $m_{\text{FW}}(X_{\phi_n})$ is constructed as

$$m_{\text{FW}}(X_{\phi_n}) = \int m_{p_{n-1} \rightarrow X_{\phi_{n-1}}}(X_{\phi_{n-1}}) m_{\text{FW}}(X_{\phi_{n-1}}) \times p(X_{\phi_n} | X_{\phi_{n-1}}) dX_{\phi_{n-1}}. \quad (11)$$

Using the equivalent circular random variable distributions, we can write

$$m_{\text{FW}}(X_{\phi_n}) = \int \mathcal{CN}(e^{j\phi_{n-1}}; \mathbf{M}_{p_{n-1} \rightarrow X_{\phi_{n-1}}}, \mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}}) \times \mathcal{CN}(e^{j\phi_{n-1}}; \mathbf{M}_{\text{FW}}^{n-1}, \mathbf{V}_{\text{FW}}^{n-1}) \times \mathcal{CN}(e^{j\Delta_{\phi_n}}; \mathbf{M}_{\Delta_{\phi_n}}, \mathbf{V}_{\Delta_{\phi}}) de^{j\phi_{n-1}}, \quad (12)$$

where $p(X_{\phi_n} | X_{\phi_{n-1}}) \propto \mathcal{CN}(e^{j\Delta_{\phi_n}}; \mathbf{M}_{\Delta_{\phi_n}}, \mathbf{V}_{\Delta_{\phi}})$ represents the circular random variable equivalent of the distribution of the phase increment process $p(\phi_n | \phi_{n-1})$, with $\mathbf{M}_{\Delta_{\phi_n}}$ is used to denote the mean of the circular random variable Wiener phase process while $\mathbf{V}_{\Delta_{\phi}}$

denotes its variance. Since the mean of the Wiener phase process is zero, we can write $p(X_{\phi_n}|X_{\phi_{n-1}})$ as $\mathcal{CN}(e^{j\phi_{n-1}}; e^{j\phi_n}, \mathbf{V}_{\Delta_\phi})$. Using the properties of the Gaussian density functions, we obtain

$$m_{\text{FW}}(X_{\phi_n}) \propto \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{\text{FW}}^n, \mathbf{V}_{\text{FW}}^n\right), \quad (13)$$

where

$$\mathbf{M}_{\text{FW}}^n = \frac{(\mathbf{M}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})(\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (\mathbf{M}_{\text{FW}}^{n-1})(\mathbf{V}_{\text{FW}}^{n-1})^{-1}}{(\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (\mathbf{V}_{\text{FW}}^{n-1})^{-1}}, \quad (14)$$

and

$$\mathbf{V}_{\text{FW}}^n = \frac{1}{(\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (\mathbf{V}_{\text{FW}}^{n-1})^{-1}} + \mathbf{V}_{\Delta_\phi}, \quad (15)$$

where \mathbf{M}_{FW}^n is the mean of the forward message and \mathbf{V}_{FW}^n is its variance.

D Message $m_{\text{BW}}(e^{j\phi_n})$

The message $m_{\text{BW}}(X_{\phi_n})$ is constructed similar to the forward message. The message can be written as

$$m_{\text{BW}}(X_{\phi_n}) \propto \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{\text{BW}}^n, \mathbf{V}_{\text{BW}}^n\right), \quad (16)$$

where

$$\mathbf{M}_{\text{BW}}^n = \frac{(\mathbf{M}_{p_{n+1} \rightarrow X_{\phi_{n+1}}})(\mathbf{V}_{p_{n+1} \rightarrow X_{\phi_{n+1}}})^{-1} + (\mathbf{M}_{\text{BW}}^{n+1})(\mathbf{V}_{\text{BW}}^{n+1})^{-1}}{(\mathbf{V}_{p_{n+1} \rightarrow X_{\phi_{n+1}}})^{-1} + (\mathbf{V}_{\text{BW}}^{n+1})^{-1}}, \quad (17)$$

and

$$\mathbf{V}_{\text{BW}}^n = \frac{1}{(\mathbf{V}_{p_{n+1} \rightarrow X_{\phi_{n+1}}})^{-1} + (\mathbf{V}_{\text{BW}}^{n+1})^{-1}} + \mathbf{V}_{\Delta_\phi}, \quad (18)$$

where \mathbf{M}_{BW}^n is the mean of the backward message and \mathbf{V}_{BW}^n is its variance.

E Message $m_{p_n \rightarrow X_n}(e^{j\angle x_n})$

The message from the phase graph to the code graph $m_{p_n \rightarrow X_n}(e^{j\angle x_n})$ can be formulated as

$$m_{p_n \rightarrow X_n}(e^{j\angle x_n}) = \int m_{X_{\phi_n} \rightarrow p_n}(e^{j\phi_n}) p(y_n|X_n, X_{\phi_n}) de^{j\phi_n}, \quad (19)$$

Using properties (A.1) and (A.2) stated in the Appendix, we can write $p(y_n|X_n, X_{\phi_n}) = \mathcal{CN}\left(e^{j\phi_n}; \frac{y_n}{e^{j\angle x_n}}, \sigma^2\right)$. Further, the message $m_{X_{\phi_n} \rightarrow p_n}(e^{j\phi_n})$ can be written as

$$\begin{aligned} m_{X_{\phi_n} \rightarrow p_n} &= \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{\text{FW}}^n, \mathbf{V}_{\text{FW}}^n\right) \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{\text{BW}}^n, \mathbf{V}_{\text{BW}}^n\right) \\ &\propto \mathcal{CN}\left(e^{j\phi_n}; \mathbf{M}_{X_{\phi_n} \rightarrow p_n}, \mathbf{V}_{X_{\phi_n} \rightarrow p_n}\right), \end{aligned} \quad (20)$$

where $\mathbf{M}_{X_{\phi_n} \rightarrow p_n}$ is the mean of the message $m_{X_{\phi_n} \rightarrow p_n}(e^{j\phi_n})$ and $\mathbf{V}_{X_{\phi_n} \rightarrow p_n}$ is the variance. These are computed using property (A.3) as

$$\mathbf{M}_{X_{\phi_n} \rightarrow p_n} = \frac{(\mathbf{M}_{\text{FW}}^n)(\mathbf{V}_{\text{FW}}^n)^{-1} + (\mathbf{M}_{\text{BW}}^n)(\mathbf{V}_{\text{BW}}^n)^{-1}}{(\mathbf{V}_{\text{FW}}^n)^{-1} + (\mathbf{V}_{\text{BW}}^n)^{-1}}, \quad (21)$$

and

$$\mathbf{V}_{X_{\phi_n} \rightarrow p_n} = \frac{1}{(\mathbf{V}_{\text{FW}}^n)^{-1} + (\mathbf{V}_{\text{BW}}^n)^{-1}}. \quad (22)$$

Therefore, using property (A.4), we evaluate (19) as

$$m_{p_n \rightarrow X_n}(e^{j\angle x_n}) \propto \mathcal{CN}\left(e^{j\angle x_n}; \mathbf{M}_{p_n \rightarrow X_n}, \mathbf{V}_{p_n \rightarrow X_n}\right), \quad (23)$$

where $\mathbf{M}_{p_n \rightarrow X_n} = \left(\frac{\mathbf{M}_{X_{\phi_n} \rightarrow p_n}}{y_n}\right)^*$ and $\mathbf{V}_{p_n \rightarrow X_n} = \frac{\sigma^2 + \mathbf{V}_{X_{\phi_n} \rightarrow p_n}}{|y_n|^2}$.

We make the following remarks. The proposed algorithm may be implemented in two ways. In the first implementation, the algorithm computes the phase noise estimates using the parameters of the messages $m_{p_n \rightarrow X_{\phi_n}}$, $m_{\text{FW}}(X_{\phi_n})$ and $m_{\text{BW}}(X_{\phi_n})$. The received data y_n is then corrected for phase noise errors. In the second implementation, the mean and variance of the message $m_{p_n \rightarrow X_n}$ are used to determine the probabilities which are sent to the code graph.

IV Results

We characterise the performance of the proposed method using simulation. We use the LDPC code of rate 0.5 [9]. A single pilot sequence is included in every 20 transmitted symbols. The phase noise is modelled as Wiener process with mean zero and variance σ_ϕ^2 . In every iteration, messages are exchanged between the phase graph and the code graph. The bit error rate (BER) that corresponds to the bit-energy-to-noise ratio $\frac{E_b}{N_o}$ is used as a performance metric. For comparison, we use the Tikhonov method [2] and a receiver where the phase noise is known. The lowest error rate is achieved using discrete messages with large number of samples. However, as discussed in [2], the computational complexity is high.

We also compare the computational complexity of the Tikhonov method and the proposed method. For a similar comparison with existing algorithms, the computational complexity is based on the exchange of symbol beliefs with the code graph. We use the total number of elementary mathematical operations and the number of memory accesses required for each iteration as a complexity metric. The Tikhonov algorithm requires 17M+17 operations and 3M+3 memory accesses. The proposed method requires 3M+51 operations and 2M+4 memory accesses for the first implementation, and 3M+56 operations and 2M+1 memory accesses for the second implementation. From this data, one can see that the proposed method achieves lower complexity especially at higher modulation orders. However, it should be mentioned that the proposed algorithm has an even lower complexity if one bases the comparison on the exchange of bit beliefs with the code graph.

Figure 2 shows the BER of the proposed method and the existing algorithms for the severe Wiener phase noise channel with $\sigma_\phi = 0.1$ rad/symbol. The LDPC code used is (3, 6) of length 4000. The curve labelled as 'Known phase' represents the case where phase

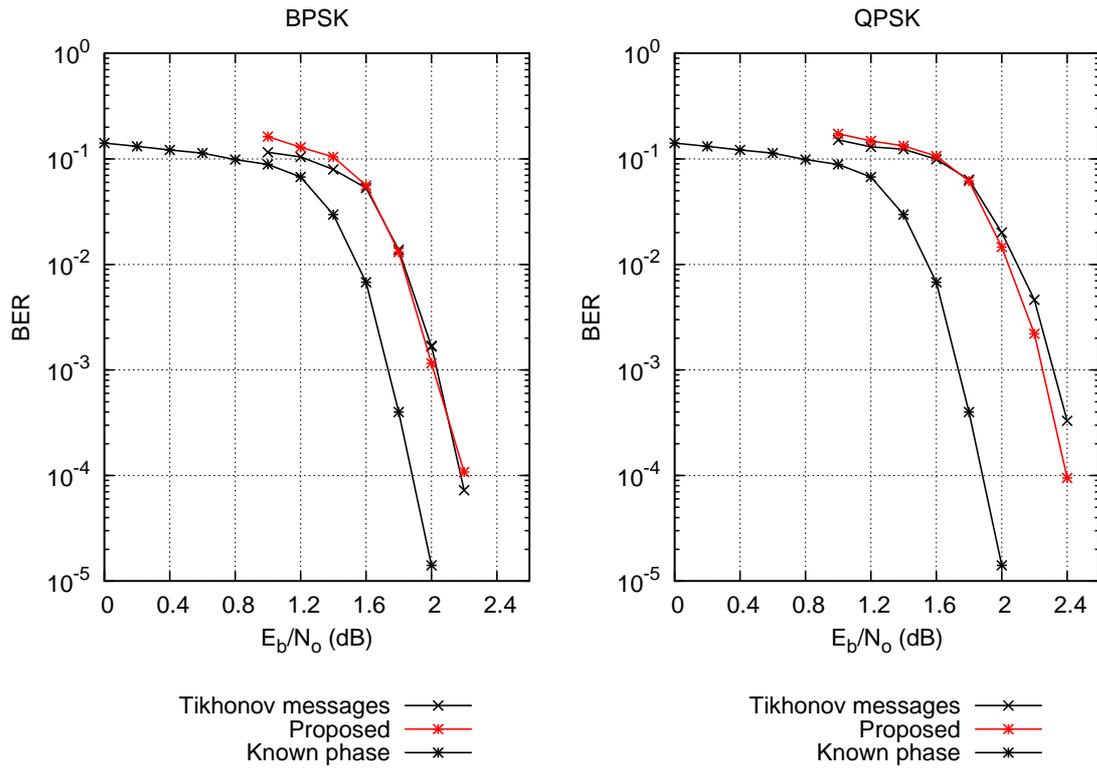


Figure 2: Bit error rate performance for BPSK and QPSK.

noise is known at the receiver. For the BPSK modulation, the proposed method using the circular random variables is shown to perform very closely to the Tikhonov method. For the QPSK modulation, the proposed method is shown to achieve similar error rates as the Tikhonov method. Figure 3 shows the BER performance of 8-PSK modulation for a code of length 2640 and for a phase noise channel with $\sigma_\phi = 0.05$ rad/symbol. The proposed algorithm is shown to effectively detect the coded data despite the higher modulation and phase noise channel. These results show that the proposed method is a competitive alternative for communications over severe phase noise channels.

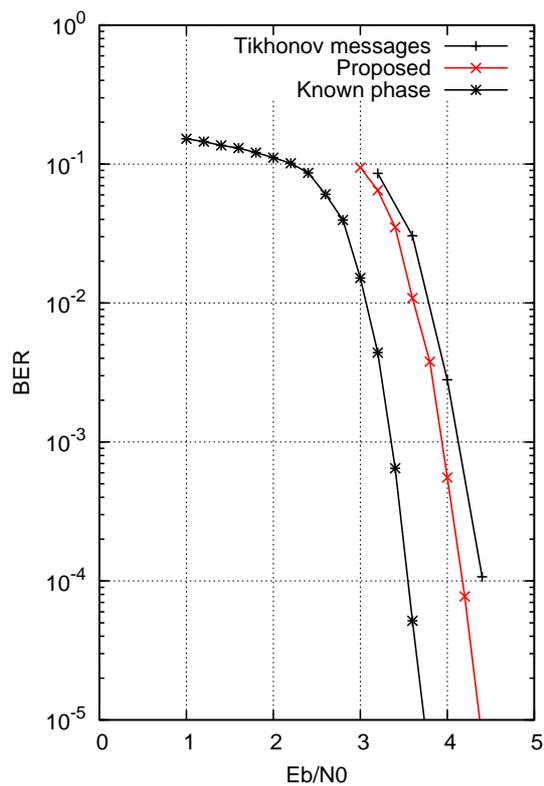


Figure 3: Bit error rate performance for 8-PSK.

V Conclusion

In this paper, we considered the iterative detection of LDPC codes in the presence of severe phase noise. We proposed a method based on the circular transformation of the random variables and presented the construction of the messages in the graphical model. The results show that the proposed method achieves similar performance as the existing algorithm for severe phase noise channels while reducing the computational complexity.

Appendix

We use properties of the normal Gaussian probability density function [10] which has been modified for complex random variables [11].

$$\mathcal{CN}(x; \mu, \sigma^2) = \mathcal{CN}(\mu; x, \sigma^2) \quad (\text{A.1})$$

$$\mathcal{CN}(ax + b; \mu, \sigma^2) = \mathcal{CN}\left(x; \frac{\mu - b}{a}, \frac{\sigma^2}{|a|^2}\right) \quad (\text{A.2})$$

$$\mathcal{CN}(x; \mu_1, \sigma_1^2)\mathcal{CN}(x; \mu_2, \sigma_2^2) \propto \mathcal{CN}\left(x; \frac{\mu_1\sigma_1^{-2} + \mu_2\sigma_2^{-2}}{\sigma_1^{-2} + \sigma_2^{-2}}, \frac{1}{\sigma_1^{-2} + \sigma_2^{-2}}\right) \quad (\text{A.3})$$

$$\int \mathcal{CN}(x; \mu_1, \sigma_1^2)\mathcal{CN}(x; \mu_2, \sigma_2^2) dx \propto \mathcal{CN}(\mu_1; \mu_2, \sigma_1^2 + \sigma_2^2) \quad (\text{A.4})$$

References

- [1] R. G. Gallager, “Low-density parity-check codes,” *IRE Transactions on Information Theory*, vol. 8, no. 1, pp. 21–28, 1962.
- [2] G. Colavolpe, A. Barbieri, and G. Caire, “Algorithms for iterative decoding in the presence of strong phase noise,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, 2005.
- [3] J. Dauwels and H.-A. Loeliger, “Phase estimation by message passing,” in *IEEE International Conference on Communications*, vol. 1, 2004, pp. 523–527.
- [4] A. Anastasopoulos and K. M. Chugg, “Adaptive iterative detection for phase tracking in turbo-coded systems,” *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2135–2144, 2001.
- [5] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [6] A. P. Worthen and W. E. Stark, “Unified design of iterative receivers using factor graphs,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 843–849, 2001.
- [7] M. Peleg, S. Shamai, and S. Galan, “Iterative decoding for coded noncoherent MPSK communications over phase-noisy AWGN channel,” *IEE Proceedings-Communications*, vol. 147, no. 2, pp. 87–95, 2000.
- [8] S. Shayovitz and D. Raphaeli, “Efficient iterative decoding of LDPC in the presence of strong phase noise,” in *International Symposium on Turbo Codes and Iterative Information Processing (ISTC)*, 2012, pp. 1–5.
- [9] D. MacKay, “D. MacKay’s Gallager code resources,” in *available at <http://www.inference.phy.cam.ac.uk/mackay/>*.
- [10] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [11] E. Aktas, “Belief propagation with Gaussian priors for pilot-assisted communication over fading ISI channels,” *IEEE Transactions on Wireless Communications*, vol. 8, no. 4, pp. 2056–2066, 2009.

Paper 3

Iterative Detection of CO-OFDM Signals with Phase Errors Using Circular Random Variables

International Journal of Communication Systems (under review)

Abstract

We propose a low-complexity graph-based iterative algorithm for the detection of coherent optical OFDM signals in the presence of residual carrier frequency offset and phase noise. We exploit the circular properties of these phase errors to derive tractable graphical messages using the Gaussian probability density function. The proposed algorithm shows similar error rate performance as the existing algorithm at a much lower implementation complexity.

I Introduction

The transmission of data with coherent optical Orthogonal Frequency Division Multiplexing (CO-OFDM) has been found to be attractive for optical networks due to its tolerance to fiber distortions such as chromatic dispersion and polarisation mode dispersion [1]. However, the longer symbol duration of the OFDM signal makes it vulnerable to phase errors caused by carrier frequency offset and phase noise. These parameters have been known to cause common phase error and intercarrier interference [2], [3]. If these effects are not estimated and compensated, the performance of the system is severely affected.

Various algorithms have been proposed to detect the CO-OFDM signal in the presence of phase errors. The existing methods can be broadly categorised as pilot based and non-pilot based approaches. In [4], pilot sequences have been inserted in the CO-OFDM symbol to estimate the common phase error. The performance improves with the number of pilot sequences at the expense of the spectral efficiency. In [5], the authors improve the spectral

efficiency using quasi-pilot sequences which can also serve as data sequences through a systematic arrangement. In [6], a radio frequency pilot tone with amplified power is inserted at the centre of the transmitted symbol and surrounded with null-subcarriers. The receiver would then use low pass filters to extract the radio frequency pilot tone and compensates for the phase errors.

There have also been other work that estimate the phase error and counteract the resulting intercarrier interference. In [7]–[9], algorithms based on the orthogonal basis expansion have been used. In [10], [11], linear interpolation based approaches have been used. In [12], a time domain approach has been proposed that is based on the sub-symbol processing. In [13], a sub-symbol based linear interpolation algorithm has been introduced. The authors in [14] use an algorithm based on the principal component estimation and elimination to lower the complexity of algorithms based on orthogonal basis expansion. Extended Kalman filtering [15] has also been proposed with small number of subcarriers. In [16], [17], algorithms based on the Gaussian basis expansion have been used which show similar complexity compared to the orthogonal basis expansion but with better linewidth tolerance.

Approaches that do not use pilot sequences have also been proposed in [18] based on the maximum likelihood algorithm. In [19], a blind phase search algorithm is presented which requires large number of test phases along with decision feedback. The work in [20], [21] have also shown estimation approaches that do not use pilot sequences. In [22], [23], the authors propose algorithms that estimate the phase noise based on a pre-defined cost function and does not require decision feedback. The use of image processing approaches have also been presented in [24]. These approaches have high implementation complexity and this scales further with increase in the size of the system.

The algorithms discussed earlier assume that there is no carrier frequency offset in the received signal. In practice however, there is some residual frequency offset error in the system which results after the dedicated estimation algorithms [25], [26] and the references therein. In [27], the authors have proposed an algorithm based on the orthogonal basis expansion to estimate and compensate for the residual frequency offset and the phase noise. The complexity of this approach is high since it involves multiple domain transformations and matrix inversions. In [26], an algorithm has been proposed that exploits the symmetry of the phase noise spectrum to estimate the residual frequency offset, whereas the phase noise has been mitigated using the principal components estimation algorithm [14]. The residual frequency offset is selected by testing multiple candidates that minimise the cost function describing the spectrum symmetry. This is further averaged over multiple symbols. Although the complexity is much lower than [27], the use of multiple test samples to estimate the frequency offset impacts the complexity of this approach. Furthermore, these algorithms require the use of pilot sequences which impact on the spectral efficiency of the system.

In this work, we detect the signal in the presence of residual carrier frequency offset and phase noise using graphical models. In [28], the authors proposed an algorithm

based on probabilistic graphical models [29] to detect the OFDM signal in the presence of frequency offset and phase noise and showed that the resulting receiver outperforms the algorithms in [30] when implemented in decision feedback system. However, its implementation complexity is high as it uses discrete samples to represent the messages of the frequency offset and phase noise. Instead of using discrete messages, we use a different approach to represent the messages in the graphical model.

The carrier frequency offset and phase noise are assumed to be continuous variables and we define these over a unit circle and exploit the circular properties. A recent study [31] has proposed a low complexity algorithm to detect low density parity check (LDPC) codes using the circular properties of the phase noise. However, the system considers the phase noise only, and the effect of the frequency offset is not shown. In this work, we use the concept of circular random variables to jointly model the phase noise and residual carrier frequency offset of the CO-OFDM system. We represent the graphical messages using complex Gaussian probability density functions, and propagate the mean and variance of the messages. We show that the proposed algorithm shows good performance and significant improvement in complexity when compared to the existing algorithm that use discrete messages. This approach can also be viewed as pilot based method for the first symbol since we use pilots at the start of each CO-OFDM frame. However, for the subsequent symbols the proposed algorithm does not require the use of pilot sequences. To the best of our knowledge there are no existing algorithms in the published work which jointly model the residual frequency offset and phase noise in CO-OFDM systems with graphical models by exploiting their circular properties.

This paper is organised as follows. Section II presents the coherent optical OFDM system and the graphical model. In Section III, the proposed method is presented and the messages are constructed using the circular random variables. Section IV presents the results and conclusions are drawn in Section V.

Notation We use bold small case and upper case letters to denote vectors and matrices. $\mathcal{CN}(x; \mu, \sigma^2)$ denotes a complex Gaussian function with mean μ and variance σ^2 . \mathbf{F} is the DFT operator with entries $\frac{1}{\sqrt{N}} \exp(-j2\pi pq/N)$ where $p, q \in [0, N - 1]$. The Hermitian transpose is denoted as $(\cdot)^H$ and the conjugate is $(\cdot)^*$.

II System Model

The transmission system using the coherent optical OFDM is shown in Figure 1. The input data bits are mapped with complex modulation sets, and then sent to the OFDM modulator which performs the inverse DFT (IDFT) operation. Thereafter, a cyclic prefix is added to mitigate the effects of the intersymbol interference. The resulting digital signal is converted to an analogue signal with the digital-to-analogue converter (DAC) and then sent to the optical modulator, where the real and imaginary parts of the electrical signal are transformed to an optical signal using laser sources. The signal is then transmitted through the optical fibers and optical amplifiers. At the coherent receiver, the optical

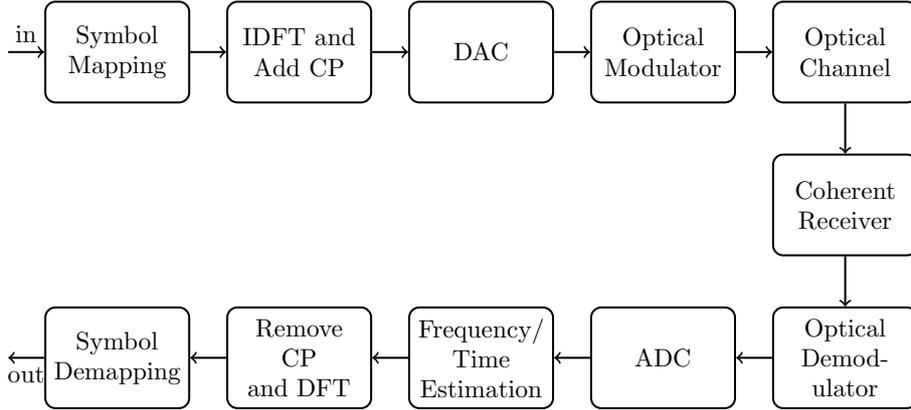


Figure 1: Coherent optical OFDM system model.

signal is converted to the electrical signal using optical demodulators, which generate the real and imaginary parts of the electrical signal. The converted signal is digitised with the analogue-to-digital converters (ADC). After estimation of the time and frequency with dedicated preambles, the receiver removes the cyclic prefix, performs the DFT operation followed by symbol demodulation to recover the transmitted data.

The received signal at time sample n can be represented as [1]

$$y_n = e^{\frac{j2\pi nv}{N}} e^{j\phi_n} [x_n * h_n] + w_n, \quad (1)$$

where x_n represents the OFDM signal, h_n is the composite channel and $*$ denotes the convolution operator. Further, v is the residual carrier frequency offset and ϕ_n is the Wiener phase noise, where $\phi_n = \phi_{n-1} + \Delta\phi$, and $\Delta\phi$ has a Gaussian distribution with zero mean and variance σ_ϕ^2 . Further, $\sigma_\phi^2 = 2\pi\beta T$ where β is the laser linewidth and T is the sampling time interval [2], and w_n represents the amplified spontaneous emission which is modelled as additive white Gaussian noise with zero-mean and variance σ^2 . In (1), the non-linear effect of the fiber transmission has been ignored since it can be compensated [32].

The received signal after ideal time estimation can be expressed as

$$\mathbf{y} = \mathbf{\Gamma}(v, \boldsymbol{\phi}) \mathbf{F}^H \mathbf{s} + \mathbf{w}, \quad (2)$$

where $\mathbf{y} = [y_0, y_1, \dots, y_{N-1}]$, $\boldsymbol{\phi} = [\phi_0, \phi_1, \dots, \phi_{N-1}]$, and $\mathbf{s} = [s_0, s_1, \dots, s_{N-1}]$ is the information data. Furthermore $\mathbf{\Gamma}$ is a diagonal matrix with entries given as $[e^{j\phi_0}, e^{j2\pi v/N + j\phi_1}, \dots, e^{j2\pi v(N-1)/N + j\phi_{N-1}}]$, and $\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]$ is the additive white Gaussian noise.

The estimation problem can be formulated as

$$\left(\hat{v}, \hat{\boldsymbol{\phi}} \right) = \underset{v, \boldsymbol{\phi}}{\operatorname{argmax}} p(v, \boldsymbol{\phi} | \mathbf{y}, \mathbf{x}), \quad (3)$$

where $p(v, \boldsymbol{\phi} | \mathbf{y}, \mathbf{x}) \propto p(\mathbf{y} | v, \boldsymbol{\phi}, \mathbf{x}) p(v) p(\boldsymbol{\phi})$. Assuming the known data symbols, $\mathbf{x} = \mathbf{F}^H \mathbf{s}$,

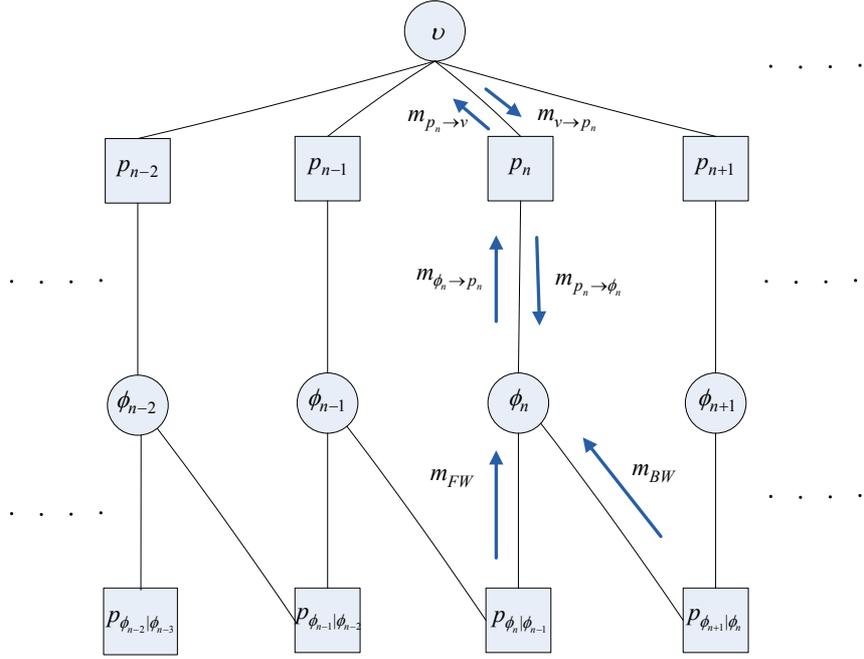


Figure 2: Graphical model.

the distribution is written as [28]

$$\begin{aligned}
 p(v, \phi | \mathbf{y}, \mathbf{x}) &\propto p(v)p(\phi_0) \prod_{n=1}^{N-1} p(\phi_n | \phi_{n-1}) \\
 &\times \prod_{n=0}^{N-1} p(y_n | v, \phi_n, x_n),
 \end{aligned} \tag{4}$$

where $p(y_n | v, \phi_n, x_n) \propto \exp \left\{ -\frac{|y_n - e^{j2\pi vn/N + j\phi_n} x_n|^2}{\sigma^2} \right\}$ and $p(\phi_n | \phi_{n-1})$ is a Gaussian function with zero mean and variance σ_ϕ^2 . A graphical model is constructed [28] for the factorised distribution in (4) and is shown in Figure 2.

III Proposed Receiver

The first step of the proposed receiver is to replace the linear random variables v and ϕ_n with the circular random variables $X_v = e^{j2\pi nv/N}$ and $X_{\phi_n} = e^{j\phi_n}$. The probability of the phase increment $p(\phi_n | \phi_{n-1})$ is also described using the circular random variable equivalent $p(X_{\phi_n} | X_{\phi_{n-1}})$. The existing graphical model shown in Figure 2 is then updated to include the circular random variables and associated factor nodes. The second step is to use the sum-product algorithm [29] to construct the graphical messages using the technique of circular random variables. This requires one to rewrite the probability distribution $p(y_n | v, \phi_n, x_n)$ as a complex Gaussian distribution $\mathcal{CN}(y_n; e^{j2\pi vn/N + j\phi_n} x_n, \sigma^2)$.

A Message $m_{p_n \rightarrow X_v}(X_v)$

The message $m_{p_n \rightarrow X_v}(X_v)$ is computed as

$$m_{p_n \rightarrow X_v}(X_v) = \int p(y_n|v, \phi_n, x_n) m_{X_{\phi_n} \rightarrow p_n}(X_{\phi_n}) dX_{\phi_n}, \quad (5)$$

and using the properties of the Gaussian density functions [33], we obtain

$$m_{p_n \rightarrow X_v}(X_v) \approx \mathcal{CN}(X_v; \mathbf{M}_{p_n \rightarrow X_v}, \mathbf{V}_{p_n \rightarrow X_v}), \quad (6)$$

where $\mathbf{M}_{p_n \rightarrow X_v} = \left(\frac{\mathbf{M}_{X_{\phi_n} \rightarrow p_n}}{\frac{y_n}{x_n}} \right)^*$ is the mean of the message and $\mathbf{V}_{p_n \rightarrow X_v} = \left(\frac{\frac{\sigma^2}{|x_n|^2} + \mathbf{V}_{X_{\phi_n} \rightarrow p_n}}{\left| \frac{y_n}{x_n} \right|^2} \right)$ is its variance.

B Message $m_{p_n \rightarrow X_{\phi_n}}(e^{j\phi_n})$

The message $m_{p_n \rightarrow X_{\phi_n}}(e^{j\phi_n})$ is formulated as

$$m_{p_n \rightarrow X_{\phi_n}} = \int p(y_n|v, \phi_n, x_n) m_{X_v \rightarrow p_n}(X_v) dX_v, \quad (7)$$

and using the properties [33], we write

$$m_{p_n \rightarrow X_{\phi_n}} \approx \mathcal{CN}(e^{j\phi_n}; \mathbf{M}_{p_n \rightarrow X_{\phi_n}}, \mathbf{V}_{p_n \rightarrow X_{\phi_n}}), \quad (8)$$

where $\mathbf{M}_{p_n \rightarrow X_{\phi_n}} = \left(\frac{\mathbf{M}_{X_v \rightarrow p_n}}{\frac{y_n}{x_n}} \right)^*$ and $\mathbf{V}_{p_n \rightarrow X_{\phi_n}} = \left(\frac{\frac{\sigma^2}{|x_n|^2} + \mathbf{V}_{X_v \rightarrow p_n}}{\left| \frac{y_n}{x_n} \right|^2} \right)$.

C Message $m_{FW}(e^{j\phi_n})$

The forward message $m_{FW}(e^{j\phi_n})$ is computed as

$$m_{FW}(X_{\phi_n}) = \int m_{p_{n-1} \rightarrow X_{\phi_{n-1}}}(X_{\phi_{n-1}}) \times m_{FW}(X_{\phi_{n-1}}) p(X_{\Delta\phi_n}) dX_{\phi_{n-1}}, \quad (9)$$

where $p(X_{\Delta\phi_n})$ is the distribution of the phase noise circular variable $X_{\Delta\phi_n} = X_{\phi_n} X_{\phi_{n-1}}^{-1}$ and can be approximated as $\mathcal{CN}(e^{j\phi_{n-1}}; e^{j\phi_n}, \sigma_\phi^2)$ [31]. Hence, we evaluate (9) as

$$m_{FW}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; \mathbf{M}_{FW}^n, \mathbf{V}_{FW}^n), \quad (10)$$

where the mean of the forward message is

$$\mathbf{M}_{FW}^n = \frac{\frac{\mathbf{M}_{p_{n-1} \rightarrow X_{\phi_{n-1}}}}{\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}}} + \frac{\mathbf{M}_{FW}^{n-1}}{\mathbf{V}_{FW}^{n-1}}}{(\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (\mathbf{V}_{FW}^{n-1})^{-1}}, \quad (11)$$

and its variance is

$$\mathbf{V}_{FW}^n = \frac{1}{(\mathbf{V}_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (\mathbf{V}_{FW}^{n-1})^{-1}} + \sigma_\phi^2. \quad (12)$$

D Message $m_{BW}(e^{j\phi_n})$

The backward message is evaluated similar to the forward message. Hence, $m_{BW}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; \mathbf{M}_{BW}^n, \mathbf{V}_{BW}^n)$. The mean and variance can be computed by substituting the variables $\mathbf{M}_{p_{n+1} \rightarrow X_{\phi_{n+1}}}$, $\mathbf{V}_{p_{n+1} \rightarrow X_{\phi_{n+1}}}$, \mathbf{M}_{BW}^{n+1} , and \mathbf{V}_{BW}^{n+1} , instead of the respective variables in (11) and (12). The details are omitted here.

E Message $m_{X_v \rightarrow p_n}(X_v)$

The message $m_{X_v \rightarrow p_n}(X_v)$ is formulated as

$$m_{X_v \rightarrow p_n}(X_v) = \prod_{k=0; k \neq n}^{N-1} m_{p_k \rightarrow X_v} \quad (13)$$

$$\approx \mathcal{CN}(X_v; \mathbf{M}_{X_v \rightarrow p_n}, \mathbf{V}_{X_v \rightarrow p_n}), \quad (14)$$

where

$$\mathbf{M}_{X_v \rightarrow p_n} = \frac{\sum_{k=0; k \neq n}^{N-1} M_{p_k \rightarrow X_v} (V_{p_k \rightarrow X_v})^{-1}}{\sum_{k=0; k \neq n}^{N-1} (V_{p_k \rightarrow X_v})^{-1}}, \quad (15)$$

and

$$\mathbf{V}_{X_v \rightarrow p_n} = \frac{1}{\sum_{k=0; k \neq n}^{N-1} (V_{p_k \rightarrow X_v})^{-1}}. \quad (16)$$

F Message $m_{X_{\phi_n} \rightarrow p_n}(X_{\phi_n})$

The message $m_{X_{\phi_n} \rightarrow p_n}(X_{\phi_n})$ is formulated as

$$m_{X_{\phi_n} \rightarrow p_n}(X_{\phi_n}) = m_{FW}(X_{\phi_n}) m_{BW}(X_{\phi_n}) \quad (17)$$

$$\approx \mathcal{CN}(e^{j\phi_n}; \mathbf{M}_{X_{\phi_n} \rightarrow p_n}, \mathbf{V}_{X_{\phi_n} \rightarrow p_n}),$$

where $\mathbf{V}_{X_{\phi_n} \rightarrow p_n} = [(\mathbf{V}_{FW}^n)^{-1} + (\mathbf{V}_{BW}^n)^{-1}]^{-1}$ and the mean is

$$\mathbf{M}_{X_{\phi_n} \rightarrow p_n} = \left(\frac{M_{FW}^n}{V_{FW}^n} + \frac{M_{BW}^n}{V_{BW}^n} \right) (\mathbf{V}_{X_{\phi_n} \rightarrow p_n}). \quad (18)$$

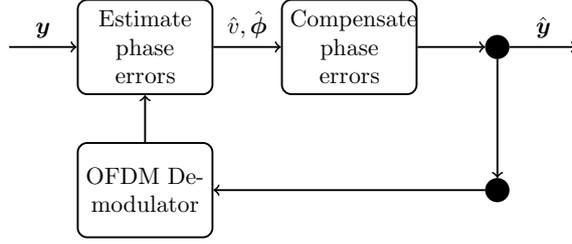


Figure 3: Estimation of phase errors with the graphical models.

G Beliefs $p(X_v)$ and $p(X_{\phi_n})$

The belief of the carrier frequency offset is computed as

$$p(X_v) = \prod_{n=0}^{N-1} m_{p_n \rightarrow X_v} \quad (19)$$

$$\approx \mathcal{CN}(X_v; \mathbf{M}_{p(X_v)}, \mathbf{V}_{p(X_v)}), \quad (20)$$

where the mean is

$$\mathbf{M}_{p(X_v)} = \frac{\sum_{n=0}^{N-1} (M_{p_n \rightarrow X_v}) (V_{p_n \rightarrow X_v})^{-1}}{\sum_{n=0}^{N-1} (V_{p_n \rightarrow X_v})^{-1}}. \quad (21)$$

Similarly, the belief of the phase noise is

$$p(X_{\phi_n}) = m_{p_n \rightarrow \phi_n}(X_{\phi_n}) m_{FW}(X_{\phi_n}) m_{BW}(X_{\phi_n}) \quad (22)$$

$$\approx \mathcal{CN}(e^{j\phi_n}; \mathbf{M}_{p(X_{\phi_n})}, \mathbf{V}_{p(X_{\phi_n})}), \quad (23)$$

where the mean is

$$\mathbf{M}_{p(X_{\phi_n})} = \frac{\frac{M_{p_n \rightarrow X_{\phi_n}}}{(V_{p_n \rightarrow X_{\phi_n}})} + \frac{M_{FW}^n}{(V_{FW}^n)} + \frac{M_{BW}^n}{(V_{BW}^n)}}{(V_{p_n \rightarrow X_{\phi_n}})^{-1} + (V_{FW}^n)^{-1} + (V_{BW}^n)^{-1}}. \quad (24)$$

The angle of (21) and (24) is used to estimate the carrier frequency offset and phase noise. The received signal sequences are then compensated with these phase error estimates as $\hat{y}_n = y_n e^{-j2\pi\hat{v}n/N} e^{-j\hat{\phi}_n}$. The transmitted bits can then be demodulated with the compensated signal sequences. The estimation of the phase errors with the proposed algorithm is summarised in Figure 3.

IV Results

We consider an OFDM symbol of 256 subcarriers, with 31 null subcarriers. A cyclic prefix of 30 samples is inserted for each symbol. The channel is assumed to be known and compensated at the receiver. The bit-error rate (BER) is used as a performance metric. For comparison, we present the performance of the proposed algorithm, the discrete algorithm

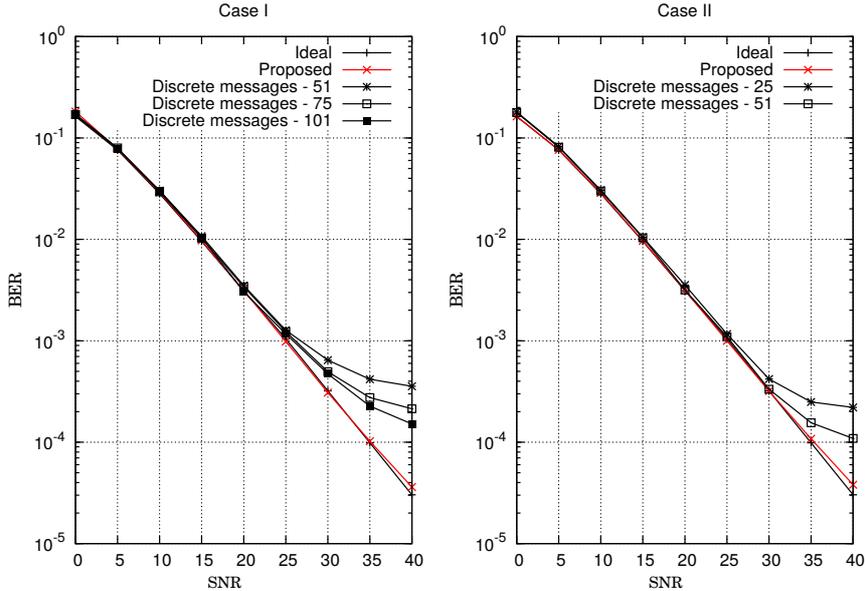


Figure 4: Error rate performance of the proposed algorithm.

[28] with equally-spaced samples of the phase errors, and an ideal algorithm with prior knowledge of the frequency offset and phase noise.

We consider two simulation cases. In the first case, the receiver has complete knowledge of the transmitted data, the residual carrier frequency offset is uniformly distributed in $(-0.2, 0.2)$ for each symbol, and the normalised laser linewidth (βT) is 2.0×10^{-4} . This case represents the initialisation of a receiver using preamble data. In the second case, the receiver operates in a decision feedback system where one known CO-OFDM symbol is inserted at the start of every 20 transmitted symbols. The residual carrier frequency offset is uniformly distributed in $(-0.05, 0.05)$ for each symbol. The normalised laser linewidth is 5.0×10^{-6} .

Figure 4 shows the uncoded BER performance of the proposed algorithm, the discrete algorithm and the ideal algorithm for Quaternary Phase Shift Keying (QPSK) modulation. The number of samples required by the discrete algorithm is also shown. For Case I, the proposed algorithm performs close to the ideal algorithm. The discrete algorithm is shown to require large message samples to approach the performance of the proposed algorithm. It would show similar performance as the proposed algorithm for sufficiently large samples. However, this would require the implementation complexity to be unnecessarily large. For Case II, the proposed algorithm is shown to outperform the discrete algorithm, and closely approaches the performance of the ideal algorithm. The discrete algorithm still requires large number of samples and seems to show error floors at high signal-to-noise ratios. These results show that the proposed algorithm is an effective algorithm in the presence of frequency offset and phase noise.

The complexity of the proposed method scales as $\mathcal{O}(iN)$, where N is the DFT size and i is the number of iterations. However, the complexity of the discrete algorithm

scales as $\mathcal{O}(iNM^2)$, where M is the number of the discrete samples that represent the messages of the phase errors required by the algorithm. The proposed and the discrete algorithms require about four iterations. Therefore, the proposed algorithm has very low complexity because it only needs to compute the mean and variance of each message in the graphical model, whereas the discrete algorithm is encumbered with processing the messages represented by large samples of the phase errors. The number of discrete samples required by the existing algorithm would be slightly improved with adaptive representation of the messages [34],[35]. However, the complexity of such an approach would still be of polynomial order. We can therefore conclude that the proposed algorithm is an attractive low complexity algorithm for CO-OFDM systems where the unknown parameters are circular random variables.

V Conclusion

We have presented a low-complexity iterative algorithm for the detection of coherent optical OFDM systems in the presence of residual carrier frequency offset and phase noise. The proposed algorithm models the carrier frequency offset and phase noise as circular random variables, and propagates the mean and variance of the Gaussian messages. The proposed algorithm has been shown to perform close to the ideal algorithm and it has low complexity.

References

- [1] W. Shieh and I. Djordjevic, *OFDM for optical communications*. Academic press, 2009.
- [2] T. Pollet, M. van Bladel, and M. Moeneclaey, “BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise,” *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 191–193, Feb./Mar./Apr. 1998.
- [3] S. Wu and Y. Bar-Ness, “A phase noise suppression algorithm for OFDM-based WLANs,” *IEEE Communications Letters*, vol. 6, no. 12, pp. 535–537, 2002.
- [4] X. Yi, W. Shieh, and Y. Tang, “Phase estimation for coherent optical OFDM,” *IEEE Photonics Technology Letters*, vol. 19, no. 12, pp. 919–921, 2007.
- [5] S. T. Le, T. Kanesan, E. Giacomidis, N. J. Doran, and A. D. Ellis, “Quasi-pilot aided phase noise estimation for coherent optical OFDM systems,” *IEEE Photonics Technology Letters*, vol. 26, no. 5, pp. 504–507, 2014.
- [6] S. Randel, S. Adhikari, and S. L. Jansen, “Analysis of RF-pilot-based phase noise compensation for coherent optical OFDM systems,” *IEEE Photonics Technology Letters*, vol. 22, no. 17, pp. 1288–1290, 2010.
- [7] C. Yang, F. Yang, and Z. Wang, “Orthogonal basis expansion-based phase noise estimation and suppression for CO-OFDM systems,” *IEEE Photonics technology letters*, vol. 22, no. 1, pp. 51–53, 2009.
- [8] —, “Phase noise suppression for coherent optical block transmission systems: A unified framework,” *Optics Express*, vol. 19, no. 18, pp. 17 013–17 020, 2011.
- [9] X. Fang, C. Yang, T. Zhang, and F. Zhang, “Orthogonal basis expansion-based phase noise suppression for PDM CO-OFDM system,” *IEEE Photonics Technology Letters*, vol. 26, no. 4, pp. 376–379, 2013.
- [10] M. E. Mousa-Pasandi and D. V. Plant, “Noniterative interpolation-based partial phase noise ICI mitigation for CO-OFDM transport systems,” *IEEE Photonics Technology Letters*, vol. 23, no. 21, pp. 1594–1596, 2011.
- [11] M. E. Mousa-Pasandi, Q. Zhuge, X. Xu, M. M. Osman, Z. A. El-Sahn, M. Chagnon, and D. V. Plant, “Experimental demonstration of non-iterative interpolation-based partial ICI compensation in 100G RGI-DP-CO-OFDM transport systems,” *Optics Express*, vol. 20, no. 14, pp. 14 825–14 832, 2012.
- [12] S. Cao, P.-Y. Kam, and C. Yu, “Time-domain blind ICI mitigation for non-constant modulus format in CO-OFDM,” *IEEE Photonics Technology Letters*, vol. 25, no. 24, pp. 2490–2493, 2013.

- [13] X. Hong, X. Hong, and S. He, “Linearly interpolated sub-symbol optical phase noise suppression in CO-OFDM system,” *Optics Express*, vol. 23, no. 4, pp. 4691–4702, 2015.
- [14] —, “Low-complexity optical phase noise suppression in CO-OFDM system using recursive principal components elimination,” *Optics Express*, vol. 23, no. 18, pp. 24 077–24 087, 2015.
- [15] T. T. Nguyen, S. T. Le, M. Wuilpart, T. Yakusheva, and P. Mégret, “Simplified extended Kalman filter phase noise estimation for CO-OFDM transmissions,” *Optics Express*, vol. 25, no. 22, pp. 27 247–27 261, 2017.
- [16] Y. Huang, Y. Chen, K. Li, Y. Han, J. Fu, J. Ma, Y. Li, J. Yu, and X. Li, “Gaussian basis expansion phase noise suppression method for CO-OFDM systems,” *Optics Express*, vol. 28, no. 17, pp. 24 343–24 352, 2020.
- [17] Y. Chen, Y. Huang, Y. Han, J. Fu, K. Li, Y. Li, and J. Yu, “A novel phase noise suppression scheme utilizing Gaussian wavelet basis expansion for PDM CO-OFDM superchannel,” *IEEE Photonics Journal*, vol. 12, no. 2, pp. 1–9, 2020.
- [18] W. Shieh, “Maximum-likelihood phase and channel estimation for coherent optical OFDM,” *IEEE Photonics technology letters*, vol. 20, no. 8, pp. 605–607, 2008.
- [19] T. Pfau, S. Hoffmann, and R. Noé, “Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for m -QAM constellations,” *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, 2009.
- [20] M. E. Mousa-Pasandi and D. V. Plant, “Data-aided adaptive weighted channel equalizer for coherent optical OFDM,” *Optics Express*, vol. 18, no. 4, pp. 3919–3927, 2010.
- [21] —, “Zero-overhead phase noise compensation via decision-directed phase equalizer for coherent optical OFDM,” *Optics Express*, vol. 18, no. 20, pp. 20 651–20 660, 2010.
- [22] S. T. Le, M. E. McCarthy, N. Mac Suibhne, P. A. Haigh, E. Giacoumidis, N. Doran, A. D. Ellis, and S. K. Turitsyn, “Decision-directed-free blind phase noise estimation for CO-OFDM,” in *Optical Fiber Communication Conference*. Optical Society of America, 2015, pp. W1E–5.
- [23] S. T. Le, P. A. Haigh, A. D. Ellis, and S. K. Turitsyn, “Blind phase noise estimation for CO-OFDM transmissions,” *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 745–753, 2015.
- [24] Z. Li, Y. Li, J. Ma, J. Chen, Y. Song, and M. Wang, “Projection histogram-assisted estimation of common phase error in coherent optical OFDM systems,” *IEEE Photonics Journal*, vol. 11, no. 3, pp. 1–9, 2019.

- [25] X. Du, T. Song, and P.-Y. Kam, “Carrier frequency offset estimation for CO-OFDM: The matched-filter approach,” *Journal of Lightwave Technology*, vol. 36, no. 14, pp. 2955–2965, 2018.
- [26] Q. Yan, X. Hong, and X. Hong, “Low-complexity residual carrier frequency offset mitigation based on spectrum symmetry for CO-OFDM systems,” *Journal of Lightwave Technology*, vol. 36, no. 22, pp. 5210–5217, 2018.
- [27] Q. Yan, X. Hong, S. He, and X. Hong, “Joint mitigation of residual carrier frequency offset and phase noise with orthogonal basis expansion in CO-OFDM,” in *Asia Communications and Photonics Conference*. Optical Society of America, 2017, pp. Su4C–2.
- [28] F. Z. Merli and G. M. Vitetta, “A factor graph approach to the iterative detection of OFDM signals in the presence of carrier frequency offset and phase noise,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 3, pp. 868–877, 2008.
- [29] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 498–519, 2001.
- [30] K. Nikitopoulos and A. Polydoros, “Phase-impairment effects and compensation algorithms for OFDM systems,” *IEEE Transactions on Communications*, vol. 53, no. 4, pp. 698–707, 2005.
- [31] S. H. Rezenom and F. Takawira, “Iterative low complexity algorithm for LDPC systems in the presence of phase noise,” *IEEE Wireless Communication Letters*, vol. 6, no. 6, pp. 794–797, 2017.
- [32] R. Dar and P. J. Winzer, “Nonlinear interference mitigation: Methods and potential gain,” *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 903–930, 2017.
- [33] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [34] A. Doucet, N. De Freitas, and N. Gordon, “An introduction to sequential monte carlo methods,” in *Sequential Monte Carlo methods in practice*. Springer, 2001, pp. 3–14.
- [35] E. B. Sudderth, A. T. Ihler, M. Isard, W. T. Freeman, and A. S. Willsky, “Nonparametric belief propagation,” *Communications of the ACM*, vol. 53, no. 10, pp. 95–103, 2010.

Paper 4

Graphical Algorithms for Optical LDPC Systems with Phase Noise Errors

Physical Communications (to be submitted)

Abstract

In this paper, we present two graphical algorithms to detect low-density parity-check (LDPC) codes for coherent optical systems with phase noise errors. The proposed algorithms are based on message propagation in graphical models. The messages to be propagated have been derived by minimising the free energy function of two combined graphical models. First, we combine the Belief Propagation (BP) and the Uniformly Reweighted Belief Propagation (URWBP) graphical models using the framework of free energy minimisation and the framework of circular random variables. We implement the algorithm using approximate and exact methods. Second, we combine the Mean Field (MF) and the URWBP graphical models using a similar framework. The results show the approximate implementation of the first algorithm outperforms the existing lowest complexity algorithm, whereas the exact implementation achieves comparable performance as the accurate and highly complex methods in the published work. Similar performance is achieved by the second algorithm.

I Introduction

The presence of phase noise has been shown to increase the error rate of communications systems. In optical communication systems [1], laser phase noise has been shown to be modelled as a Wiener process [2], [3]. The effect of phase noise channels on the information capacity has also been shown in [4]–[6]. This adverse effect on the error rate and information capacity may be decreased, but not avoided, with high precision oscillators at

the transmitter and receiver. This would increase the cost of the communications system. Furthermore, the optical channel inherently introduces phase noise into the communication system. Since phase noise can not be avoided, it is necessary that effective receivers be designed to detect data in the presence of phase noise.

The estimation of coded data in the presence of phase noise for optical systems can be classified as graph-based, log-likelihood ratio based and code-based algorithms. The graph-based algorithms jointly estimate the phase noise and the code with messages that propagate between the code graph and the phase graph. The log-likelihood ratio based algorithms modify existing metrics to include information about the phase noise [7]–[9]. These algorithms have shown improvements compared to existing approaches. The code-based algorithms construct specialised codes to withstand the effects of the Gaussian noise and the phase noise [10]–[13]. One category of these algorithms uses differential encoding along with an existing code to counteract the phase noise channel. Another category [11], [13] constructs dedicated codes that can be combined with the differential encoding. The construction of such dedicated codes is an intensive and complex process which requires the use of protographs and other approaches. The graph-based algorithms are attractive over the other approaches since the construction of specialised codes is not required. Furthermore, the graphical algorithms have low complexity compared to the other algorithms.

There is a growing interest in the use of graphical models to estimate phase noise in communication systems. Graphical models have been widely used in the decoding of capacity-approaching codes. Since the work of Gallager which used graphical models to show the relationship among coded bits [14], there has been significant research that highlighted the relationship between graphical models and probabilistic decoding. The belief propagation (BP) algorithm [15] has been successfully used in the decoding of capacity-approaching codes such as the Low Density Parity Check (LDPC) codes and turbo codes [16], [17]. This success has sparked significant research interest as it has demonstrated low-complexity decoding of capacity approaching error-correcting codes. The belief propagation algorithm was found to give exact results in tree graphs and usually performed well in cyclic graphs [18] although strict guarantees for this performance could not always be established. Further research effort revealed better understanding of the behaviour of the belief propagation algorithm.

The work in [19] has established a relationship between the belief propagation algorithm and the free energy function commonly found in statistical mechanics. The Bethe free energy is one such system of defining the energy content of a physical system [19]. It has been demonstrated that the fixed points of the loopy belief propagation algorithm are equivalent to the stationary points of the constrained Bethe free energy function of the graphical structure. Other interesting properties that relate to the stationary points of the constrained Bethe free energy function and the fixed point equations of the belief propagation algorithm are studied in [19], [20]. Furthermore, the work of [19] also describes a method to construct alternate message expressions using the regions of a graphical model.

Riegler *et al.* [21] exploited this principle of region selection and obtained message passing algorithms that combine the mean field (MF) [22] and belief propagation messages, and demonstrated its use in the joint estimation of channel parameters and decoding of an Orthogonal Frequency Division Multiplexing (OFDM) system. This combined MF/BP algorithm has been used in diverse set of problems [23], [24], where it has been shown to outperform the standard BP algorithm.

The authors in [25] have proposed assigning probabilistic weights to the edges of a graph by minimising the upper bound of the log-partition function. In [26], this approach has been simplified to use uniform weights, resulting in an algorithm denoted as Uniformly Re-weighted Belief Propagation (URWBP), and has been shown to outperform the BP algorithm.

In this work, we propose iterative algorithms that combine existing graphical models using the framework of free-energy minimisation and the framework of phase noise representation as a circular random variable. We apply the proposed algorithms to detect LDPC signals in the presence of phase noise. Several methods have addressed this problem where the main aspect is to represent the phase noise belief in the graphical model with an existing function, and subsequently propagate its parameters. In [27], the phase noise belief is represented with the Tikhonov probability density function. In [28], the phase noise belief is represented as a cluster of Tikhonov probability density functions. In [29], the phase noise is represented as a circular random variable and its belief is approximated with the complex Gaussian density function. For coded systems with moderate phase noise variance, the Tikhonov algorithm [27] shows good performance, and the circular random variable algorithm [29] achieves comparable performance with lower complexity. For severe phase noise variance, the Tikhonov cluster algorithm [28] and the algorithm in [30] achieve good performance at the cost of high implementation complexity. This work emphasises on systems with high modulation schemes and moderate phase noise variance, which is too complex for the method in [28].

The URWBP algorithm has been used for LDPC decoding in [26] for systems with no phase noise. For systems with phase noise, the existing graphical algorithms for joint estimation and decoding use the standard belief propagation to construct the messages [27]–[29]. In contrast, the proposed algorithms use messages constructed from the combination of graphical models. Therefore, the main aspect of this paper is two fold. First, we combine the BP and the URWBP graphical models using the concept of free energy, and implement the optimised algorithm using the framework of circular random variables. We implement the optimised algorithms using approximate and exact methods. The results show that these methods outperform the existing non-optimised algorithm in the published work [29]. Second, we combine the MF and the URWBP graphical models using a similar framework. We show that the combined MF and BP is the minimum free energy of the combined graphical model, and achieves comparable error rates as existing methods. The emphasis of this work is on systems with moderate phase noise variance, which are too complex for the method in [28].

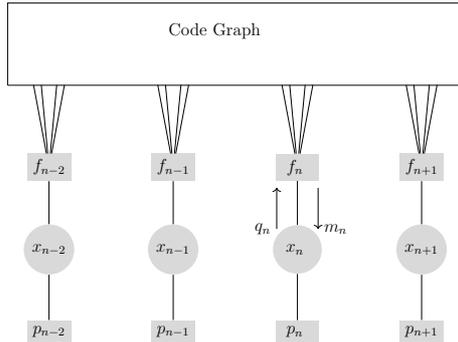


Figure 1: Basic description of a coded system.

The paper is organised as follows. Section II presents the background. Section III formulates the detection problem of LDPC codes in Wiener phase noise channels. Section IV presents the message construction using the proposed algorithms. Section VI discusses the results and Section VII presents the conclusion.

Notation. Scalar random variables are denoted using lower case letters, while bold lower case letters denote vectors. The expression $\mathcal{CN}(x; \mu, \sigma^2)$ denotes a complex Gaussian function with mean μ and variance σ^2 . The conjugate operator is denoted as $(\cdot)^*$.

II Background

This section presents the background on the several phase noise estimation approaches for the optical LDPC systems. As stated in the introduction, these approaches can be broadly classified as log-likelihood ratio based approaches, code-based approaches, and graph-based approaches.

Assume a sequence of bit streams b_n has been LDPC coded and modulated to form a complex sequence a_n . The received signal is $y_n = x_n + w_n$, where w_n represents the noise sequence. Figure 1 represents a basic description of this system, where the p_n are factor nodes as a function of x_n and w_n , and f_n is the mapping function for a 16-QAM scheme. The block shown as the “Code Graph” governs the LDPC code structure. The symbols q_n and m_n represent the respective messages propagated to and from the code graph. The message propagated from f_n to the code graph can be viewed as the existing log-likelihood ratios. The existing computation of the log-likelihood ratios take into account the received signal, the modulation symbols, and the variance of the Gaussian noise.

Now, assume that phase noise ϕ_n is present in the system, which would modify the received signal as $y_n = x_n e^{j\phi_n} + w_n$. Existing approaches would still compute the log-likelihood ratios as described earlier. However, there have been some work that modify the log-likelihood metric to include the statistical parameters of the phase noise. In [7], this metric is formulated by performing numerical methods on the expectation which includes the phase noise statistics with sufficient number of samples. In [9], the authors formulate the log-likelihood ratio with knowledge of the distribution of the Gaussian phase noise.

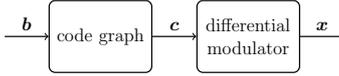


Figure 2: Description of a code-based transmitter.

In [8], similar formulation of the log-likelihood ratio is exploited for the coded optical OFDM systems, with an LDPC code of short length. Although such approaches are scarce in the published work, some improvement of the bit-error rate has been reported.

The code-based approaches can be categorised as methods that use an existing code followed with a differential modulator, and methods that construct dedicated codes for the phase noise channel [11], [13] which is also followed with a differential modulator. The first approach would be easier to implement since it uses an existing code, whereas the second approach would be complex to implement since it requires the construction of specialised codes to encounter the phase noise channel. Since existing codes are commonly constructed for the Gaussian noise channel, these specialised constructions that also take into account the phase noise channel have shown performance improvements. Figure 2 summarises the transmission process of the code-based approaches. The transmitted bits \mathbf{b} are encoded with an existing code or the specially constructed code. The encoded bits \mathbf{c} are then sent to the differential modulator which enables better handling of the phase noise. This results in the complex sequence \mathbf{x} which is sent through the Gaussian and the phase noise channels. At the receiver, the process is reversed to obtain the transmitted bits.

The graph-based approaches exploit the structure of the statistical properties of the phase noise. This statistical representation is captured and included in the graphical model of the estimation process. Figure 3 shows the modification of the system with the inclusion of the statistical graphical description of the phase noise, which we denote as “Phase Graph”. The messages are then propagated between the code graph and the phase graph until convergence is achieved. The messages propagated from the code graph are depicted as $m_{x_n \rightarrow p_n}$, whereas the message that propagate from the phase graph are depicted as $m_{p_n \rightarrow x_n}$. This approach is followed in this paper.

III System Model

We consider the detection of coded optical systems in the presence of phase noise. The binary information signal b_n is encoded and modulated using symbols from a set \mathcal{A} of size M to form a complex symbol x_n . The received signal is

$$y_n = x_n e^{j\phi_n} + w_n, \quad n = 0 \dots N - 1 \quad (1)$$

where ϕ_n is a Wiener phase noise random variable generated as $\phi_n = \phi_{n-1} + \Delta\phi$, where $\Delta\phi$ has a Gaussian distribution with zero mean and variance $\sigma_\phi^2 = 2\pi\beta T_s$. Further, β denotes the laser linewidth and T_s denotes the sampling time interval. The additive white

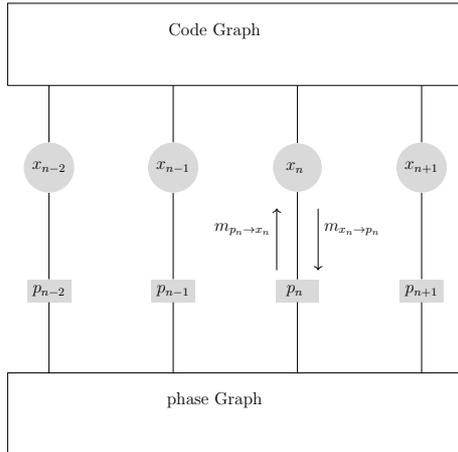


Figure 3: Systemic description of graph-based algorithm.

Gaussian noise w_n has zero mean and variance $E[|w_n|^2] = \sigma^2$.

The aim is to estimate \mathbf{x} and $\boldsymbol{\phi}$ that minimises the joint probability distribution

$$p(\mathbf{x}, \boldsymbol{\phi} | \mathbf{y}) \propto p(\mathbf{y} | \boldsymbol{\phi}, \mathbf{x}) p(\boldsymbol{\phi}) p(\mathbf{x}), \quad (2)$$

where $p(\mathbf{y} | \boldsymbol{\phi}, \mathbf{x}) \propto \prod_{i=0}^{N-1} p(y_n | x_n, \phi_n)$ and $p(y_n | x_n, \phi_n) \propto \exp\left(\frac{-|y_n - x_n e^{j\phi_n}|^2}{\sigma^2}\right)$. The distribution of the phase noise is $p(\boldsymbol{\phi}) \propto p(\phi_0) \prod_{i=1}^{N-1} p(\phi_n | \phi_{n-1})$, where $p(\phi_0)$ is a uniform distribution of the initial phase noise sample ϕ_0 and $p(\phi_n | \phi_{n-1})$ is a Gaussian function with zero mean and variance σ_ϕ^2 .

Figure 4 shows the graphical model that describes the estimation problem [27], [28] in (2) where the phase noise samples are replaced with the circular random variables $X_{\phi_n} = e^{j\phi_n}$, and the distribution of the successive phase noise samples is represented as $p(X_{\phi_n} | X_{\phi_{n-1}})$. Further, p_n represents the distribution $p(y_n | x_n, \phi_n)$. The figure also shows the messages that propagate in the graphical model.

IV Proposed Algorithms

In this section, we present the proposed algorithms which are based on the minimisation of the free energy of the graphical model and on the representation of the phase noise as a circular random variable. For the first algorithm, we combine the belief propagation and the uniformly reweighted belief propagation, and refer to it as the *combined BP and URWBP* algorithm, For the second algorithm, we combine the mean field and the uniformly reweighted belief propagation, and refer to it as the *combined MF and URWBP* algorithm.

In Appendix A, the derivation of a generic reweighted belief propagation algorithm is presented, where the messages are summarised using expressions (A.9), (A.11), and (A.12). These expressions are used to formulate the messages of the combined BP and URWBP algorithm. In Appendix B, the derivation of the combined MF and a generic reweighted

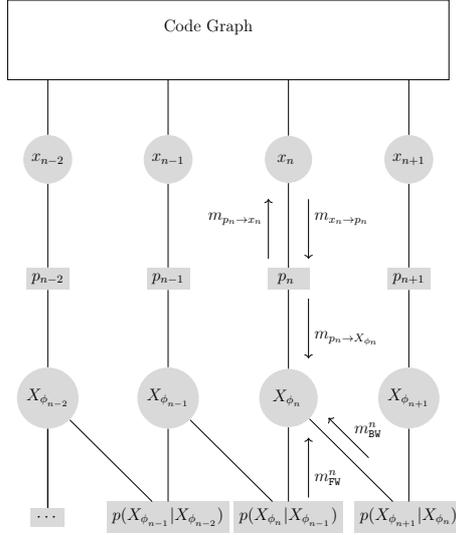


Figure 4: Graphical model of (2).

BP algorithm is presented, where the messages are constructed using expressions (B.4), (B.6), (B.7), (B.8), and (B.9). These expressions are used to formulate the messages of the combined MF and BP, and the combined MF and URWBP algorithms. In the following sections, we present the message construction of the proposed algorithms on the graph shown in Figure 4.

A Combined BP and URWBP Algorithm

This algorithm uses the BP graph to model the factor node $p(y_n|x_n, \phi_n)$ and its neighbouring variable nodes x_n and X_{ϕ_n} , and the URWBP graph to model the factor node $p(X_{\phi_n}|X_{\phi_{n-1}})$ and its neighbouring variable nodes X_{ϕ_n} and $X_{\phi_{n-1}}$. The message construction is based on [26], where for the BP factor nodes, a reweighting factor of $\rho_a = 1$ is used. For the URWBP factor nodes, a uniform reweighting factor of $\rho_a = \rho$ is used. The messages are constructed as follows.

Constructing the message $m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})$

The message $m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})$ is formulated as (A.11)

$$m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n}) = \int m_{x_n \rightarrow p_n}(x_n) p(y_n|x_n, X_{\phi_n}) dx_n, \quad (3)$$

where $m_{x_n \rightarrow p_n}(x_n) \approx \mathcal{CN}(x_n; M_{x_n \rightarrow p_n}, V_{x_n \rightarrow p_n})$ has mean $M_{x_n \rightarrow p_n}$ and variance $V_{x_n \rightarrow p_n}$. Further, we write $p(y_n|x_n, X_{\phi_n}) = \mathcal{CN}(x_n; \frac{y_n}{e^{j\phi_n}}, \sigma^2)$ and evaluate (3) as [29]

$$m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{p_n \rightarrow X_{\phi_n}}, V_{p_n \rightarrow X_{\phi_n}}), \quad (4)$$

where $M_{p_n \rightarrow X_{\phi_n}} = \left(\frac{y_n}{M_{x_n \rightarrow p_n}} \right)$ and $V_{p_n \rightarrow X_{\phi_n}} = \left(\frac{V_{x_n \rightarrow p_n} + \sigma^2}{|M_{x_n \rightarrow p_n}|^2} \right)$.

Constructing the forward message $m_{\text{FW}}(X_{\phi_n})$

The forward message is formulated (A.11) as

$$m_{\text{FW}}(X_{\phi_n}) = \int \frac{p(X_{\phi_{n-1}})}{m_{\text{BW}}(X_{\phi_{n-1}})} (p(X_{\phi_n} | X_{\phi_{n-1}}))^{\frac{1}{\rho}} dX_{\phi_{n-1}}, \quad (5)$$

where the belief $p(X_{\phi_{n-1}})$ is formulated (A.9) as

$$p(X_{\phi_{n-1}}) = m_{p_{n-1} \rightarrow X_{\phi_{n-1}}}(X_{\phi_{n-1}}) \times (m_{\text{FW}}(X_{\phi_{n-1}}))^{\rho} (m_{\text{BW}}(X_{\phi_{n-1}}))^{\rho}. \quad (6)$$

We use the complex Gaussian representation of messages $m_{p_{n-1} \rightarrow X_{\phi_{n-1}}}(X_{\phi_{n-1}})$, $m_{\text{FW}}(X_{\phi_{n-1}})$, and $m_{\text{BW}}(X_{\phi_{n-1}})$ with respective mean and variances of $M_{p_{n-1} \rightarrow X_{\phi_{n-1}}}$, $V_{p_{n-1} \rightarrow X_{\phi_{n-1}}}$, M_{FW}^{n-1} , V_{FW}^{n-1} , M_{BW}^{n-1} and V_{BW}^{n-1} . Further, the distribution of the phase noise increment is also represented using the complex Gaussian function as $p(X_{\phi_n} | X_{\phi_{n-1}}) \approx \mathcal{CN}(e^{j\phi_{n-1}}; e^{j\phi_n}, \sigma_{\phi}^2)$. We can then write the forward message as a Gaussian function

$$m_{\text{FW}}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{\text{FW}}^n, V_{\text{FW}}^n), \quad (7)$$

with variance

$$V_{\text{FW}}^n = A_{\text{FW}}^n + \rho \sigma_{\phi}^2, \quad (8)$$

where

$$A_{\text{FW}}^n = \left(\frac{1}{V_{p_{n-1} \rightarrow X_{\phi_{n-1}}}} + \frac{\rho}{V_{\text{FW}}^{n-1}} + \frac{\rho - 1}{V_{\text{BW}}^{n-1}} \right)^{-1}, \quad (9)$$

and with mean

$$M_{\text{FW}}^n = ((M_{p_{n-1} \rightarrow X_{\phi_{n-1}}})(V_{p_{n-1} \rightarrow X_{\phi_{n-1}}})^{-1} + (M_{\text{FW}}^{n-1})(V_{\text{FW}}^{n-1}/\rho)^{-1} + (M_{\text{BW}}^{n-1})(V_{\text{BW}}^{n-1}/\rho - 1)^{-1})(A_{\text{FW}}^n). \quad (10)$$

Constructing the backward message $m_{\text{BW}}(X_{\phi_n})$

The backward message is formulated using (A.11)

$$m_{\text{BW}}(X_{\phi_n}) = \int \frac{p(X_{\phi_{n+1}})}{m_{\text{FW}}(X_{\phi_{n+1}})} (p(X_{\phi_{n+1}} | X_{\phi_n}))^{\frac{1}{\rho}} dX_{\phi_{n+1}}, \quad (11)$$

where $p(X_{\phi_{n+1}})$, $m_{\text{FW}}(X_{\phi_{n+1}})$, and $p(X_{\phi_{n+1}} | X_{\phi_n})$ represent the distributions at sample $n + 1$. Since the derivation of backward message is similar to the forward message, we use the complex Gaussian representation of messages $m_{p_{n+1} \rightarrow X_{\phi_{n+1}}}(X_{\phi_{n+1}})$, $m_{\text{FW}}(X_{\phi_{n+1}})$, and

$m_{\text{BW}}(X_{\phi_{n+1}})$ with respective mean and variances of $M_{p_{n+1} \rightarrow X_{\phi_{n+1}}}$, $V_{p_{n+1} \rightarrow X_{\phi_{n+1}}}$, M_{FW}^{n+1} , V_{FW}^{n+1} , M_{BW}^{n+1} and V_{BW}^{n+1} . We express the message as $m_{\text{BW}}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{\text{BW}}^n, V_{\text{BW}}^n)$ with variance

$$V_{\text{BW}}^n = A_{\text{BW}}^n + \rho\sigma_\phi^2, \quad (12)$$

where

$$A_{\text{BW}}^n = \left(\frac{1}{V_{p_{n+1} \rightarrow X_{\phi_{n+1}}}} + \frac{\rho}{V_{\text{BW}}^{n+1}} + \frac{\rho-1}{V_{\text{FW}}^{n+1}} \right)^{-1}, \quad (13)$$

and mean

$$\begin{aligned} M_{\text{BW}}^n &= ((M_{p_{n+1} \rightarrow X_{\phi_{n+1}}})(V_{p_{n+1} \rightarrow X_{\phi_{n+1}}})^{-1} \\ &\quad + (M_{\text{BW}}^{n+1})(V_{\text{BW}}^{n+1}/\rho)^{-1} \\ &\quad + (M_{\text{FW}}^{n+1})(V_{\text{FW}}^{n+1}/\rho-1)^{-1})(A_{\text{BW}}^n). \end{aligned} \quad (14)$$

Constructing the phase noise belief $p(X_{\phi_n})$

The phase noise belief is formulated as (A.9)

$$p(X_{\phi_n}) = m_{p_n \rightarrow X_{\phi_n}}(X_{\phi_n})(m_{\text{FW}}(X_{\phi_n}))^\rho(m_{\text{BW}}(X_{\phi_n}))^\rho, \quad (15)$$

and is computed as $p(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{X_{\phi_n}}, V_{X_{\phi_n}})$, where $V_{X_{\phi_n}}$ denotes the variance of the phase noise and is computed as

$$V_{X_{\phi_n}} = \frac{1}{(V_{p_n \rightarrow X_{\phi_n}})^{-1} + (V_{\text{FW}}^n/\rho)^{-1} + (V_{\text{BW}}^n/\rho)^{-1}}, \quad (16)$$

and $M_{X_{\phi_n}}$ denotes the mean of the phase noise and is computed as

$$\begin{aligned} M_{X_{\phi_n}} &= ((M_{p_n \rightarrow X_{\phi_n}})(V_{p_n \rightarrow X_{\phi_n}})^{-1} + (M_{\text{FW}}^n)(V_{\text{FW}}^n/\rho)^{-1} \\ &\quad + (M_{\text{BW}}^n)(V_{\text{BW}}^n/\rho)^{-1})(V_{X_{\phi_n}}). \end{aligned} \quad (17)$$

The estimate of the phase noise is then computed from the argument $\angle M_{X_{\phi_n}}$. The message to the code graph is obtained with standard demapping approaches which rotate the received signal with the phase noise estimates.

B Combined MF and URWBP Algorithm

In this section, we present the message construction of the combined MF and URWBP algorithm. The MF algorithm models the factor node $p(y_n|x_n, X_{\phi_n})$ and its neighbouring variables nodes, while the URWBP algorithm models the factor node $p(X_{\phi_n}, X_{\phi_{n-1}})$ and its neighbouring variable nodes. The message construction is based on [21], [26], where we use the URWBP algorithm instead of the BP algorithm. Hence, the URWBP factor nodes

are modelled using a uniform reweighting factor $\rho_a = \rho$. The combined MF and BP is a special formulation where $\rho = 1$. The messages are constructed as follows.

Constructing the message $m_{p_n \rightarrow X_{\phi_n}}^{\text{MF}}(X_{\phi_n})$

The mean field message $m_{p_n \rightarrow X_{\phi_n}}^{\text{MF}}(X_{\phi_n})$ is formulated as (B.8)

$$m_{p_n \rightarrow X_{\phi_n}}^{\text{MF}}(X_{\phi_n}) = \exp \left(\sum_{x_n \in \mathcal{A}} p(x_n) \log p(y_n | x_n, X_{\phi_n}) \right), \quad (18)$$

where $p(x_n)$ is the belief of the symbol x_n . Therefore

$$\begin{aligned} m_{p_n \rightarrow \phi_n}^{\text{MF}}(X_{\phi_n}) &\propto \exp \left(- \sum_{x_n \in \mathcal{A}} p(x_n) \frac{|y_n - x_n e^{j\phi_n}|^2}{\sigma^2} \right) \\ &\propto \mathcal{CN}(e^{j\phi_n}; M_{p_n \rightarrow \phi_n}^{\text{MF}}, V_{p_n \rightarrow \phi_n}^{\text{MF}}), \end{aligned} \quad (19)$$

where $M_{p_n \rightarrow X_{\phi_n}}^{\text{MF}} = \frac{E[|x_n|^2]}{y_n^* E[x_n]}$ and $V_{p_n \rightarrow X_{\phi_n}}^{\text{MF}} = \frac{\sigma^2 E[|x_n|^2]}{|y_n|^2 |E[x_n]|^2}$. $E[\cdot]$ denotes the expectation operator.

Constructing the forward message $m_{\text{FW}}(X_{\phi_n})$

The forward message is constructed as (B.6)

$$m_{\text{FW}}(X_{\phi_n}) = \int \frac{p(X_{\phi_{n-1}})}{m_{\text{BW}}(X_{\phi_{n-1}})} (p(X_{\phi_n} | X_{\phi_{n-1}}))^{\frac{1}{\rho}} dX_{\phi_{n-1}}, \quad (20)$$

where we use (B.4) to write

$$\begin{aligned} p(X_{\phi_{n-1}}) &= m_{p_{n-1} \rightarrow X_{\phi_{n-1}}}^{\text{MF}}(X_{\phi_{n-1}}) \\ &\quad \times (m_{\text{FW}}(X_{\phi_{n-1}}))^{\rho} (m_{\text{BW}}(X_{\phi_{n-1}}))^{\rho}. \end{aligned} \quad (21)$$

We then write the forward message as $m_{\text{FW}}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{\text{FW}}^n, V_{\text{FW}}^n)$. The mean and variance are similar in form to (10) and (8), and are not included here to avoid repetition.

Constructing the backward message $m_{\text{BW}}(X_{\phi_n})$

The backward messages are constructed in a similar method as the forward messages and can be evaluated to give a Gaussian function $m_{\text{BW}}(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{\text{BW}}^n, V_{\text{BW}}^n)$ with mean and variance that are similar in form to (14) and (12).

Constructing the phase noise belief $p(X_{\phi_n})$

The phase noise belief is formulated as

$$p(X_{\phi_n}) = m_{p_n \rightarrow X_{\phi_n}}^{\text{MF}}(X_{\phi_n}) (m_{\text{FW}}(X_{\phi_n}))^{\rho} (m_{\text{BW}}(X_{\phi_n}))^{\rho}, \quad (22)$$

and is computed as $p(X_{\phi_n}) \approx \mathcal{CN}(e^{j\phi_n}; M_{X_{\phi_n}}, V_{X_{\phi_n}})$, where $V_{X_{\phi_n}}$ denotes the variance of the phase noise and is computed as

$$V_{X_{\phi_n}} = \frac{1}{(V_{p_n \rightarrow X_{\phi_n}}^{\text{MF}})^{-1} + (V_{\text{FW}}^n/\rho)^{-1} + (V_{\text{BW}}^n/\rho)^{-1}}, \quad (23)$$

and $M_{X_{\phi_n}}$ denotes the mean of the phase noise and is computed as

$$M_{X_{\phi_n}} = ((M_{p_n \rightarrow X_{\phi_n}}^{\text{MF}})(V_{p_n \rightarrow X_{\phi_n}})^{-1} + (M_{\text{FW}}^n)(V_{\text{FW}}^n/\rho)^{-1} + (M_{\text{BW}}^n)(V_{\text{BW}}^n/\rho)^{-1})(V_{X_{\phi_n}}). \quad (24)$$

The argument of the mean is used as the phase noise estimate. The message to the code graph is computed as discussed in the previous subsection.

V Complexity

The complexity of the proposed algorithms is presented relative to the existing circular random variable algorithm [29], where the number of extra multiplication operators is used as metric. Assume the variable a represents the number of extra multiplication operators required to compute the third term of the mean of the forward and backward messages. This requires two extra scalar-complex multiplications and six scalar-scalar multiplications. The complexity of the approximate Proposed Algorithm I with optimal ρ would require about $\mathcal{O}(a)$ extra multiplication operations, whereas the exact Proposed Algorithm I with optimal ρ and the Proposed Algorithm II would require about $\mathcal{O}(6M + a)$ extra multiplication operations, where M is the size of the modulation scheme. The complexity of the algorithm in [28] is extremely high in comparison to the proposed algorithms, and is only used here because of its best performance.

VI Results

We characterise the performance of the proposed algorithms to detect optical LDPC systems in Wiener phase noise channels. The LDPC code is of length 2640 and rate 0.5 [31]. A single known symbol is inserted every 20 transmitted symbols to aid the estimation process. The sampling rate is 25 GS/s and the normalised laser linewidth (βT_s) is 0.0004. The bit-error rate (BER) is evaluated at each optical signal-to-noise ratio (OSNR) [32]. We compare the performance of the proposed algorithms with the circular random variable algorithm [29], the Tikhonov cluster algorithm with a divergence measure of $\epsilon = 15$ and unlimited components [28], and the algorithm with no phase noise.

In this section, the combined BP and URWBP algorithm is referred to as ‘‘Proposed Algorithm I’’ whereas the combined MF and URWBP algorithm is referred to as ‘‘Proposed Algorithm II’’. Further, we represent the message from the code graph using exact and approximate methods. In the exact method, we compute the mean and variance of the

message using the complete modulation symbols. In the approximate method, we use only one symbol to compute these parameters.

Figure 5 shows the effect of the reweighting factors on the error rate of the proposed algorithms for 16-QAM at OSNR of 10.6 dB. The error rate of the exact and approximate methods of the Proposed Algorithm I is shown to be minimised at $\rho \approx 0.9$. This shows that the existing BP only algorithm is not the minimum free energy point of the graphical model, and the performance could be optimised further. The Proposed Algorithm II shows minimum error rate at $\rho = 1$, and the error rate increases for low ρ . This shows that the Proposed Algorithm II has minimum free energy at $\rho = 1$, which is the combined MF and BP algorithm.

Figure 6 shows the BER performance of the proposed algorithms for 16-QAM. The approximate Proposed Algorithm I with $\rho = 0.89$ is shown to outperform the circular random variable algorithm. It has to be stated the circular random variable algorithm is equivalent to the approximate Proposed Algorithm I with $\rho = 1$. The results also show that the optimised approximate Proposed Algorithm I improves the performance of the circular random variable algorithm with minimal increase in complexity. This performance is shown to be comparable to the accurate and computationally complex Tikhonov cluster algorithm. Furthermore, the exact Proposed Algorithm I with $\rho = 0.9$ shows close performance compared to the Tikhonov cluster algorithm. This shows there is minimal performance gain in using the exact method for moderate phase noise variance. The results of the Proposed Algorithm II show similar performance as the optimised approximate Proposed Algorithm I and the exact Proposed Algorithm I. This is because the algorithm also uses an exact method to compute the mean and variance of the messages from the code graph.

Figure 7 shows the performance of the proposed algorithms as a variation of the normalised laser linewidth (βT_s) for OSNR of 10.6 dB. The approximate Proposed Algorithm I with $\rho = 0.89$ is shown to improve the performance of the non-optimised algorithm. Further, the exact Proposed Algorithm I with $\rho = 0.9$ shows similar performance as the Tikhonov cluster algorithm. The exact Proposed Algorithm II with $\rho = 1.0$ also shows similar performance for moderate values of the phase noise variance. This also shows the good balance between performance and complexity of the optimised proposed algorithm I in comparison to the other algorithms.

In summary, the results show optimising graphical models improves performance. In the broader sense, the proposed algorithms show alternative options of achieving comparable performance to the best algorithms in the published work for moderate phase noise channels. In view of complexity and performance, the approximate proposed algorithm improves the performance of existing low complexity methods at a slight implementation cost. From the point of view of free energy, the optimised BP graph has lower free energy compared to the optimised MF graph, at least for the graphical model of the phase noise estimation problem of coded systems.

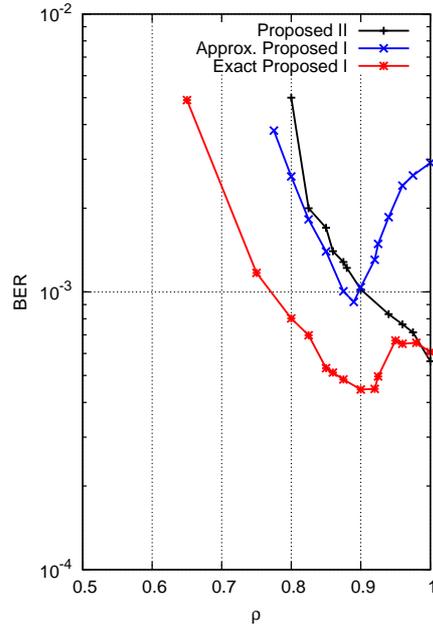


Figure 5: BER of the proposed algorithms for varying ρ .

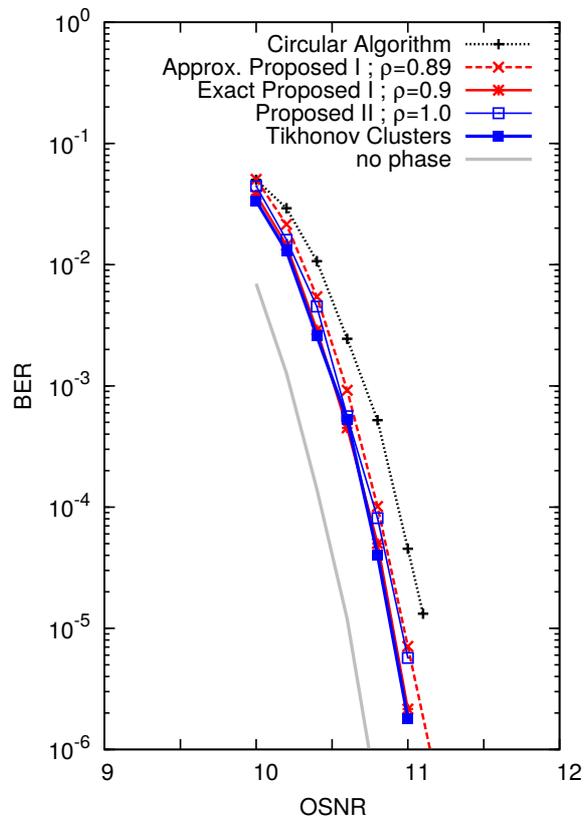


Figure 6: BER of the proposed algorithms.

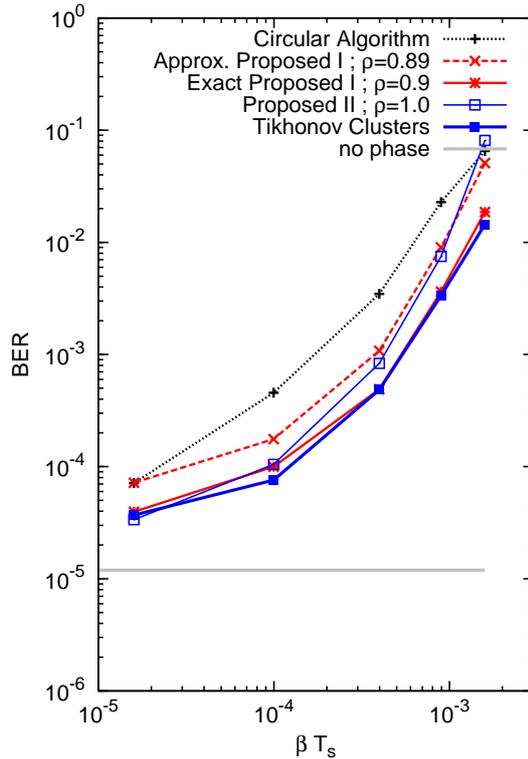


Figure 7: BER as a variation of the normalised laser linewidth.

VII Conclusion

In this paper, we have proposed two combined graphical algorithms to detect LDPC codes for coherent optical systems with phase noise errors. These algorithms are based on the minimisation of the free energy of the graphical model combined with an implementation of the phase noise as a circular random variable. The first proposed algorithm exploits messages obtained from the combined BP and URWBP algorithms. The second proposed algorithm exploits messages obtained from the combined URWBP and MF algorithms. These algorithms have been implemented with the exact and approximate approaches. The first proposed algorithm with the approximate implementation has shown improved performance compared to an existing low complexity algorithm with slight increase in complexity. The exact implementation has shown improved performance compared to the non-optimised algorithm of similar complexity. Further, the use of circular random variables have maintained the complexity advantage of existing algorithms. In general, the results have shown the advantages of constructing messages derived from the minimised free energy function of graphical models.

Appendix A

Derivation of a Reweighted BP Algorithm

Notation A graphical model \mathcal{G} is represented by the set of variable nodes \mathcal{V} , a set of factor nodes \mathcal{F} and a set of edges connecting these nodes \mathcal{E} . The mini-graph G_a is the smallest graph constructed according to [19], and consists of a single factor node and a set of variable nodes \mathbf{x}_{G_a} . The graph G_{BP} represents a graph composed entirely of BP mini-graphs. The graph G_{MF} represents a graph composed entirely of MF mini-graphs. The mini-graph G_a has a reweighting factor ρ_a , where $0 < \rho_a \leq 1$. A value of $\rho_a = 1$ indicates a BP graph.

The free energy of a reweighted belief propagation graph can be formulated as

$$\begin{aligned}
F = & - \sum_{G_a \in G_{BP}} \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log f_{G_a}(\mathbf{x}_{G_a}) \\
& + \sum_{G_a \in G_{BP}} \rho_a \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log b_{G_a}(\mathbf{x}_{G_a}) \\
& + \sum_{i \in \mathcal{V}} \sum_{x_i} b_i(x_i) \log b_i(x_i) \\
& - \sum_{G_a \in G_{BP}} \rho_a \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log \prod_{i \in \mathcal{N}(a)} b_i(x_i), \tag{A.1}
\end{aligned}$$

where $b_{G_a}(\mathbf{x}_{G_a})$ denotes the belief of the mini-graph G_a and $b_i(x_i)$ denotes the belief of each variable node i . The constraints are

$$\sum_{x_i} b_i(x_i) - 1 = 0, \quad \forall i \in \mathcal{V}, \tag{A.2}$$

$$\sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) - 1 = 0, \quad \forall G_a \in G_{BP}, \tag{A.3}$$

$$\sum_{\mathbf{x}_{G_a \sim x_i}} b_{G_a}(\mathbf{x}_{G_a}) - b_i(x_i) = 0, \quad \forall G_a \in G_{BP}, \tag{A.4}$$

$$b_i(x_i) \geq 0, \tag{A.5}$$

$$b_{G_a}(\mathbf{x}_{G_a}) \geq 0, \quad \forall G_a \in G_{BP}. \tag{A.6}$$

The Lagrangian \mathcal{L} can be written as

$$\begin{aligned}
\mathcal{L} = & F + \sum_{i \in \mathcal{V}} \lambda_i \left(\sum_{x_i} b_i(x_i) - 1 \right) \\
& + \sum_{G_a \in G_{BP}} \gamma_a \left(\sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) - 1 \right) \\
& + \sum_{(ai) \in \mathcal{E}_{BP}} \eta_{ai}(x_i) \left(\sum_{\mathbf{x}_{G_a \sim x_i}} b_{G_a}(\mathbf{x}_{G_a}) - b_i(x_i) \right), \tag{A.7}
\end{aligned}$$

where λ_i , γ_a and η_{ai} are the respective Lagrange multipliers of the constraint functions. The belief $b_i(x_i)$ is obtained by setting the $\frac{\partial \mathcal{L}}{\partial b_i(x_i)} = 0$ as

$$b_i(x_i) = \exp \left(\sum_{(ai) \in \mathcal{E}_{BP}} \eta_{ai}(x_i) \right). \quad (\text{A.8})$$

If we set $\eta_{ai}(x_i) = \log(m_{a \rightarrow i}(x_i))^{\rho_a}$, we obtain

$$b_i(x_i) = \prod_{a \in \mathcal{N}(i)} (m_{a \rightarrow i}(x_i))^{\rho_a}. \quad (\text{A.9})$$

Similarly, the belief $b_{G_a}(\mathbf{x}_{G_a})$ is obtained by setting $\frac{\partial \mathcal{L}}{\partial b_{G_a}(\mathbf{x}_{G_a})} = 0$ as

$$b_{G_a}(\mathbf{x}_{G_a}) = (f_{G_a}(\mathbf{x}_{G_a}))^{1/\rho_a} \prod_{i \in \mathcal{N}(a)} \frac{b_i(x_i)}{m_{a \rightarrow i}(x_i)}. \quad (\text{A.10})$$

The message passing rules can be obtained by using the consistency constraints $b_i(x_i) = \sum_{\mathbf{x}_{G_a} \sim x_i} b_{G_a}(\mathbf{x}_{G_a})$. Hence, we obtain

$$m_{a \rightarrow i}(x_i) = \sum_{\mathbf{x}_{G_a} \sim x_i} (f_{G_a}(\mathbf{x}_{G_a}))^{1/\rho_a} \prod_{j \in \mathcal{N}(a) \sim i} \frac{b_j(x_j)}{m_{a \rightarrow j}(x_j)}. \quad (\text{A.11})$$

Further, we can set $m_{j \rightarrow a}(x_j) = \frac{b_j(x_j)}{m_{a \rightarrow j}(x_j)}$ and evaluate as

$$m_{j \rightarrow a}(x_j) = (m_{a \rightarrow j}(x_j))^{\rho_a - 1} \prod_{b \in \mathcal{N}(j) \sim a} (m_{b \rightarrow j}(x_j))^{\rho_b}. \quad (\text{A.12})$$

Appendix B

Derivation of the Combined MF and Reweighted BP Algorithm

The free energy function of the combined MF and reweighted BP graph can be formulated as

$$\begin{aligned}
F = & - \sum_{G_a \in G_{BP}} \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log f_{G_a}(\mathbf{x}_{G_a}) \\
& - \sum_{G_a \in G_{MF}} \sum_{\mathbf{x}_{G_a}} \prod_{x_i \in \mathbf{x}_{G_a}} b_i(x_i) \log f_{G_a}(\mathbf{x}_{G_a}) \\
& + \sum_{G_a \in G_{BP}} \rho_a \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log b_{G_a}(\mathbf{x}_{G_a}) \\
& + \sum_{i \in \mathcal{V}} \sum_{x_i} b_i(x_i) \log b_i(x_i) \\
& - \sum_{G_a \in G_{BP}} \rho_a \sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) \log \prod_{x_i \in \mathbf{x}_{G_a}} b_i(x_i). \tag{B.1}
\end{aligned}$$

The Lagrangian can be written as

$$\begin{aligned}
\mathcal{L} = & F + \sum_{i \in \mathcal{V}} \lambda_i \left(\sum_{x_i} b_i(x_i) - 1 \right) \\
& + \sum_{G_a \in G_{BP}} \gamma_a \left(\sum_{\mathbf{x}_{G_a}} b_{G_a}(\mathbf{x}_{G_a}) - 1 \right) \\
& + \sum_{(ai) \in \mathcal{E}_{BP}} \eta_{ai}(x_i) \left(\sum_{\mathbf{x}_{G_a} \sim x_i} b_{G_a}(\mathbf{x}_{G_a}) - b_i(x_i) \right), \tag{B.2}
\end{aligned}$$

where λ_i , γ_a and η_{ai} are the respective Lagrangian multipliers for the variable nodes, factor nodes and edges of the graphical model. The belief of each variable node is obtained by $\frac{\partial \mathcal{L}}{\partial b_i(x_i)} = 0$ as

$$\begin{aligned}
b_i(x_i) = & \exp \sum_{a \in \mathcal{N}_{MF}(i)} \underbrace{\prod_{j \in \mathcal{N}(a) \sim i} b_j(x_j) \log f_{G_a}(\mathbf{x}_{G_a})}_{\log m_{a \rightarrow i}(x_i)} \\
& + \sum_{a \in \mathcal{N}_{BP}(i)} \frac{\eta_{ai}(x_i)}{\log(m_{a \rightarrow i}(x_i))^{\rho_a}}. \tag{B.3}
\end{aligned}$$

Using the definition $\log m_{a \rightarrow i}(x_i) = \prod_{j \in \mathcal{N}(a) \sim i} b_j(x_j) \log f_{G_a}(\mathbf{x}_{G_a})$ and $\log(m_{a \rightarrow i}(x_i))^{\rho_a} = \eta_{ai}(x_i)$, the belief $b_i(x_i)$ is further evaluated as

$$b_i(x_i) = \prod_{a \in \mathcal{N}_{MF}(i)} m_{a \rightarrow i}(x_i) \prod_{a \in \mathcal{N}_{BP}(i)} \log(m_{a \rightarrow i}(x_i))^{\rho_a}. \tag{B.4}$$

The belief $b_{G_a}(\mathbf{x}_{G_a})$ is similarly obtained using $\frac{\partial \mathcal{L}}{\partial b_{G_a}(\mathbf{x}_{G_a})} = 0$ as

$$b_{G_a}(\mathbf{x}_{G_a}) = (f_{G_a}(\mathbf{x}_{G_a}))^{1/\rho_a} \prod_{i \in \mathcal{N}(a)} \frac{b_i(x_i)}{m_{a \rightarrow i}(x_i)}. \quad (\text{B.5})$$

Using the constraint $\sum_{\mathbf{x}_{G_a \sim x_i}} b_{G_a}(\mathbf{x}_{G_a}) = b_i(x_i)$, we obtain the message from the URWBP factor node, i.e., $a \in G_{BP}$, as

$$m_{a \rightarrow i}(x_i) = \sum_{\mathbf{x}_{G_a \sim x_i}} (f_{G_a}(\mathbf{x}_{G_a}))^{1/\rho_a} \prod_{j \in \mathcal{N}(a) \sim i} \frac{b_j(x_j)}{m_{a \rightarrow j}(x_j)}. \quad (\text{B.6})$$

The message from a variable node to the URWBP factor node can be evaluated as

$$m_{j \rightarrow a}(x_j) = (m_{a \rightarrow j}(x_j))^{\rho_a - 1} \prod_{b \in \mathcal{N}_{MF}(j) \sim a} m_{b \rightarrow j}(x_j) \prod_{b \in \mathcal{N}_{BP}(j) \sim a} (m_{b \rightarrow j}(x_j))^{\rho_b}. \quad (\text{B.7})$$

In order to obtain the messages from a MF factor node, i.e., $a \in G_{MF}$, we use

$$m_{a \rightarrow i}(x_i) = \exp \left(\sum_{\mathbf{x}_{G_a \sim x_i}} \prod_{j \in \mathcal{N}(a) \sim i} b_j(x_j) \log f_{G_a}(\mathbf{x}_{G_a}) \right). \quad (\text{B.8})$$

Similarly, we can write the expressions for a message from a variable node to a MF factor node, i.e., $a \in G_{MF}$

$$m_{j \rightarrow a}(x_j) = \prod_{b \in \mathcal{N}_{MF}(j) \sim a} m_{b \rightarrow j}(x_j) \prod_{b \in \mathcal{N}_{BP}(j) \sim a} (m_{b \rightarrow j}(x_j))^{\rho_b}. \quad (\text{B.9})$$

References

- [1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, “Capacity limits of optical fiber networks,” *Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, 2010.
- [2] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli, “Empirical modeling and simulation of phase noise in long-haul coherent optical transmission systems,” *Optics Express*, vol. 19, no. 23, pp. 22 455–22 461, 2011.
- [3] G. J. Foschini and G. Vannucci, “Characterizing filtered light waves corrupted by phase noise,” *IEEE Transactions on Information Theory*, vol. 34, no. 6, pp. 1437–1448, 1988.
- [4] L. Barletta, M. Magarini, and A. Spalvieri, “The information rate transferred through the discrete-time Wiener’s phase noise channel,” *Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1480–1486, 2012.
- [5] L. Barletta, M. Magarini, S. Pecorino, and A. Spalvieri, “Upper and lower bounds to the information rate transferred through first-order Markov channels with free-running continuous state,” *IEEE Transactions on Information Theory*, vol. 60, no. 7, pp. 3834–3844, 2014.
- [6] A. Barbieri and G. Colavolpe, “On the information rate and repeat-accumulate code design for phase noise channels,” *IEEE Transactions on Communications*, vol. 59, no. 12, pp. 3223–3228, 2011.
- [7] I. B. Djordjevic and T. Wang, “On the LDPC-coded modulation for ultra-high-speed optical transport in the presence of phase noise,” in *IEEE Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, 2013, pp. 1–3.
- [8] S. Cao, P.-Y. Kam, C. Yu, and X. Cheng, “Pilot-tone assisted log-likelihood ratio for LDPC coded CO-OFDM system,” *IEEE Photonics Technology Letters*, vol. 26, no. 15, pp. 1577–1580, 2014.
- [9] T. Koike-Akino, D. S. Millar, K. Kojima, and K. Parsons, “Phase noise-robust LLR calculation with linear/bilinear transform for LDPC-coded coherent communications,” in *CLEO: Science and Innovations*. Optical Society of America, 2015, pp. SW1M–3.
- [10] F. Yu, N. Stojanovic, F. Hauske, D. Chang, Z. Xiao, G. Bauch, D. Pflueger, C. Xie, Y. Zhao, L. Jin *et al.*, “Soft-decision LDPC turbo decoding for DQPSK modulation in coherent optical receivers,” in *European Conference and Exposition on Optical Communications*. Optical Society of America, 2011, pp. We–10.

- [11] B. Matuz, G. Liva, E. Paolini, M. Chiani, and G. Bauch, “Low-rate non-binary LDPC codes for coherent and blockwise non-coherent AWGN channels,” *IEEE Transactions on Communications*, vol. 61, no. 10, pp. 4096–4107, 2013.
- [12] S. Bellini, M. Ferrari, A. Tomasoni, C. Costantini, L. Razzetti, and G. Gavioli, “LDPC design for block differential modulation in optical communications,” *IEEE Journal of Lightwave Technology*, vol. 33, no. 1, pp. 78–88, 2015.
- [13] T. Ninacs, B. Matuz, G. Liva, and G. Colavolpe, “Short non-binary low-density parity-check codes for phase noise channels,” *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 4575–4584, 2019.
- [14] R. Gallager, “Low-density parity-check codes,” *MIT Press*, 1963.
- [15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [16] D. J. MacKay and R. M. Neal, “Near Shannon limit performance of low density parity check codes,” *Electronics letters*, vol. 32, no. 18, pp. 1645–1646, 1996.
- [17] R. J. McEliece, D. J. C. MacKay, and J.-F. Cheng, “Turbo decoding as an instance of Pearl’s “belief propagation” algorithm,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 2, pp. 140–152, 1998.
- [18] K. P. Murphy, Y. Weiss, and M. I. Jordan, “Loopy belief propagation for approximate inference: An empirical study,” in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1999, pp. 467–475.
- [19] J. S. Yedidia, W. T. Freeman, and Y. Weiss, “Constructing free-energy approximations and generalized belief propagation algorithms,” *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2282–2312, 2005.
- [20] T. Heskes *et al.*, “Stable fixed points of loopy belief propagation are minima of the Bethe free energy,” *Advances in neural information processing systems*, vol. 15, pp. 359–366, 2003.
- [21] E. Riegler, G. E. Korkelund, C. N. Manchón, M.-A. Badiu, and B. H. Fleury, “Merging belief propagation and the mean field approximation: A free energy approach,” *IEEE Transactions on Information Theory*, vol. 59, no. 1, pp. 588–602, 2013.
- [22] M. J. Wainwright, M. I. Jordan *et al.*, “Graphical models, exponential families, and variational inference,” *Foundations and Trends in Machine Learning*, vol. 1, no. 1–2, pp. 1–305, 2008.

- [23] Z. Yuan, C. Zhang, Z. Wang, Q. Guo, S. Wu, and X. Wang, “A low-complexity receiver using combined BP–MF for joint channel estimation and decoding in OFDM systems,” *arXiv preprint*, 2016.
- [24] D. J. Jakubisin, R. M. Buehrer, and C. R. da Silva, “Probabilistic receiver architecture combining BP, MF, and EP for multi-signal detection,” *arXiv preprint arXiv:1604.04834*, 2016.
- [25] M. Wainwright, T. Jaakkola, and A. Willsky, “A new class of upper bounds on the log partition function,” *IEEE Transactions on Information Theory*, vol. 51, no. 7, pp. 2313–2335, 2005.
- [26] H. Wymeersch, F. Penna, and V. Savic, “Uniformly reweighted belief propagation for estimation and detection in wireless networks,” *Wireless Communications, IEEE Transactions on*, vol. 11, no. 4, pp. 1587–1595, 2012.
- [27] G. Colavolpe, A. Barbieri, and G. Caire, “Algorithms for iterative decoding in the presence of strong phase noise,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 9, pp. 1748–1757, 2005.
- [28] S. Shayovitz and D. Raphaeli, “Message passing algorithms for phase noise tracking using Tikhonov mixtures,” *IEEE Transactions on Communications*, vol. 64, no. 1, pp. 387–401, 2016.
- [29] S. H. Rezenom and F. Takawira, “Iterative low complexity algorithm for LDPC systems in the presence of phase noise,” *IEEE Wireless Communication Letters*, vol. 6, no. 6, pp. 794–797, 2017.
- [30] A. Kreimer and D. Raphaeli, “Efficient low-complexity phase noise resistant iterative joint phase estimation and decoding algorithm,” *IEEE Transactions on Communications*, vol. 66, no. 9, pp. 4199–4210, 2018.
- [31] D. MacKay, “D. Mackay’s Gallager code resources,” in *available at <http://www.inference.phy.cam.ac.uk/mackay/>*.
- [32] A. Leven and L. Schmalen, “Status and recent advances on forward error correction technologies for lightwave systems,” *Journal of lightwave technology*, vol. 32, no. 16, pp. 2735–2750, 2014.

Paper 5

Parameter Based Computation of Information Bounds for Wiener Phase Noise Channels

IEEE Communications Letters (to be submitted)

Abstract

In this paper, we propose parameter based methods to compute the mutual information bounds of the Wiener phase noise channel. The proposed methods represent the phase noise distributions as Gaussian probability density functions of circular random variables, and compute the information bounds using parameters of the Gaussian function. The results are accurate at low-to-moderate signal-to-noise ratios, and closely approach the exact mutual information at high signal-to-noise ratios. Furthermore, the complexity of the proposed methods is significantly lower than the most accurate approach of computing the mutual information.

I Introduction

There is a significant interest to develop methods that increase the information transferred through an optical fiber. The general characteristics of the optical fiber have been discussed in [1]. One aspect of the optical fiber is the presence of phase noise which results from the inaccuracies of laser sources and other non-linear sources. It has been shown that the effect of the laser linewidth can be modelled as a Wiener phase noise process [2], [3], which is similar to the modelling of an oscillator phase noise in wired and wireless systems. Since phase noise has been known to negatively affect a communication system, this has motivated efforts that characterise the performance of the optical fiber channel in the presence of phase noise.

The authors in [4] propose simulation-based methods to compute the upper and lower bounds of the mutual information. These methods have been used in subsequent works to compute the mutual information bounds of the Wiener phase noise channel. In [5], a method is proposed to compute the information lower bound of the Wiener phase noise channel by taking discrete samples of the phase noise. In [6],[7], the discrete phase method is proposed to compute the information upper bound. In [8], the information lower bound is computed with phase noise distributions formed with particle filters. The accuracy of these discrete phase methods improves with the number of phase noise samples, which leads to high implementation complexity. However, there are few existing low complexity approaches of computing the mutual information bounds. In [9], a method based on the Kalman filter is proposed to compute the information lower bound. In [10], the information upper bound is computed by approximating the phase noise distributions with the Tikhonov probability density function and the exponential Fourier series. In [11], the information lower bound is computed by approximating the phase noise distribution with multiple Tikhonov probability density functions for systems with low phase noise variance and large modulation sets.

In this work, we emphasise on parameter based methods to compute the information bounds of the Wiener phase noise channel. In [12], a framework of circular random variables is proposed for the iterative detection of coded systems in Wiener phase noise systems. Hence, we exploit this framework and propose parameter based methods to compute the mutual information bounds of the Wiener phase noise channel by approximating the phase noise distributions as Gaussian probability density functions. The results are shown to be accurate for practical values of the information rate. To compute the mutual information bounds, the proposed methods only require the parameters of the Gaussian function. Therefore, the complexity of the proposed methods is significantly lower than the discrete phase method which has the most accurate information bounds at the cost of very high complexity.

This paper is organised as follows. Section II formulates the problem of computing the mutual information bounds. Section III presents the proposed parameter based method to compute the mutual information lower bound. Section IV presents the proposed parameter based method to compute the mutual information upper bound. The results are presented in Section V and the conclusion is discussed in Section VI.

Notation The probability density function of a complex Gaussian random variable x with mean μ and variance σ^2 is denoted as $\mathcal{CN}(x; \mu, \sigma^2)$. The vector (x_1, x_2, \dots, x_n) is denoted as x_1^n , or simply as X . For a given probability density function $p(X)$, its information measure is denoted as $H(X)$. Further, $\overline{H}(X)$ denotes an upper bound on $H(X)$ while $\underline{H}(X)$ denotes a lower bound.

II Problem Formulation

The system model is described as $y_n = x_n e^{j\theta_n} + w_n$ for $n = 1, \dots, M$, where each input sequence x_n is selected from a constrained complex modulation set \mathcal{A} with a uniform distribution. The Wiener phase noise sequence is generated as $\theta_n = \theta_{n-1} + \Delta$, where Δ is a Gaussian parameter with zero mean and variance σ_θ^2 . The additive white Gaussian noise sequence w_n has zero mean and variance $E[|w_n|^2] = \sigma^2$. The mutual information between the input sequences $X = [x_1, \dots, x_M]$ and output sequences $Y = [y_1, \dots, y_M]$ is computed as

$$I(X; Y) = \sum_{X, Y} p(X, Y) \log_2 \frac{p(X, Y)}{p(X)p(Y)}, \quad (1)$$

where $p(X, Y)$ denotes the joint distribution, whereas $p(X)$ and $p(Y)$ denote the respective distributions of X and Y . Since the expression in (1) could not be computed precisely, one would have to compute its upper and lower bounds. Therefore, the problem is to establish the mutual information bounds as $\underline{I}(X; Y) \leq I(X; Y) \leq \bar{I}(X; Y)$ where $\underline{I}(X; Y)$ denotes the mutual information lower bound and $\bar{I}(X; Y)$ denotes the upper bound. These bounds are computed using a general framework which we briefly discuss. The lower bound can be formulated as [4]

$$\underline{I}(X; Y) = \mathbb{E}[\log_2 \frac{q(Y|X)}{q(Y)}] \quad (2)$$

$$\approx \frac{1}{n} \log_2 q(y_1^n | x_1^n) - \frac{1}{n} \log_2 q(y_1^n), \quad (3)$$

where $\mathbb{E}[\cdot]$ denotes the expectation operator with respect to $p(X, Y)$, and the approximation follows from the asymptotic equipartition property [13] for sufficiently large number of sequences. The distribution $q(Y|X)$ is an auxiliary channel [4] that approximates the exact distribution $p(Y|X)$, and $q(Y) = \int q(Y|X)p(X) dX$ is a distribution that approximates $p(Y)$. The upper bound can be formulated as [4]

$$\bar{I}(X; Y) = \mathbb{E}[\log_2 \frac{p(Y|X)}{q(Y)}] \quad (4)$$

$$\approx \frac{1}{n} \log_2 p(y_1^n | x_1^n) - \frac{1}{n} \log_2 q(y_1^n). \quad (5)$$

In contrast to the lower bound, the computation of the upper bound requires exact knowledge of the channel model $p(Y|X)$.

III Proposed Lower Bound

The mutual information lower bound is computed with the approximate distributions of the terms shown in (2). We approximate these distributions using the framework of circular random variables and compute their bounds.

A Compute $q(Y|X)$

We express the distribution $p(y_1^n|x_1^n)$ as

$$p(y_1^n|x_1^n) = \int p(y_1^n|x_1^n, e^{j\theta_n}) de^{j\theta_n}, \quad (6)$$

where

$$\begin{aligned} p(y_1^n|x_1^n, e^{j\theta_n}) &= \int p(y_1^{n-1}|x_1^{n-1}, e^{j\theta_{n-1}}) \\ &\quad \times p(y_n|x_n, e^{j\theta_n})p(e^{j\theta_n}|e^{j\theta_{n-1}}) de^{j\theta_{n-1}}. \end{aligned} \quad (7)$$

Hence, we approximate with circular random variables as

$$\begin{aligned} q(y_1^n|x_1^n, e^{j\theta_n}) &\approx \int \mathcal{CN}(e^{j\theta_{n-1}}; \dot{M}_{n-1}, \dot{V}_{n-1}) \\ &\quad \times \mathcal{CN}(y_n; x_n e^{j\theta_n}, \sigma^2) \mathcal{CN}(e^{j\theta_n}; e^{j\theta_{n-1}}, \sigma_\theta^2) de^{j\theta_{n-1}} \end{aligned} \quad (8)$$

$$= a_n \mathcal{CN}(e^{j\theta_n}; \dot{M}_n, \dot{V}_n), \quad (9)$$

where the properties of Gaussian distributions have been exploited to obtain (9), where

$$a_n = \mathcal{CN}(y_n; x_n \dot{M}_{n-1}, \sigma^2 + |x_n|^2 \dot{V}_{n-1} + |x_n|^2 \sigma_\theta^2). \quad (10)$$

The mean is computed as

$$\dot{M}_n = \frac{(\frac{y_n}{x_n})(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{M}_{n-1})(\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}{(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}, \quad (11)$$

and the variance as

$$\dot{V}_n = \frac{1}{(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}. \quad (12)$$

Therefore, we compute the bound on $q(y_1^n|x_1^n)$ by averaging over $\log_2 \frac{1}{a_n}$.

B Compute $q(Y)$

This is computed by approximating $q(y_1^n)$ as

$$q(y_1^n) \approx \int q(y_1^n|e^{j\theta_n}) de^{j\theta_n}, \quad (13)$$

where $q(y_1^n | e^{j\theta_n})$ is approximated from (8) as

$$q(y_1^n | e^{j\theta_n}) \approx \sum_{x_n \in \mathcal{A}} p(x_n) \int \mathcal{CN}(e^{j\theta_{n-1}}; \dot{M}_{n-1}, \dot{V}_{n-1}) \times \mathcal{CN}(e^{j\theta_{n-1}}; e^{j\theta_n}, \sigma_\theta^2) \mathcal{CN}(y_n; x_n e^{j\theta_n}, \sigma^2) de^{j\theta_{n-1}} \quad (14)$$

$$\stackrel{(a)}{=} \sum_{x_n \in \mathcal{A}} \alpha_n(x_n) \mathcal{CN}(e^{j\theta_n}; \dot{M}_n(x_n), \dot{V}_n(x_n)), \quad (15)$$

where (a) follows from the properties of Gaussian distributions [14], and

$$\alpha_n(x_n) = p(x_n) \mathcal{CN}(x_n \dot{M}_{n-1}; y_n, \sigma^2 + |x_n|^2 \dot{V}_{n-1} + |x_n|^2 \sigma_\theta^2). \quad (16)$$

Further, $\dot{M}_n(x_n)$ describes the mean which is calculated as

$$\dot{M}_n(x_n) = \frac{(\frac{y_n}{x_n})(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{M}_{n-1})(\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}{(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}, \quad (17)$$

and $\dot{V}_n(x_n)$ describes the variance computed as

$$\dot{V}_n(x_n) = \frac{1}{(\frac{\sigma^2}{|x_n|^2})^{-1} + (\dot{V}_{n-1} + \sigma_\theta^2)^{-1}}. \quad (18)$$

Therefore, we compute the bound on $q(y_1^n)$ by averaging over $\log_2 \frac{1}{\sum_{x_n \in \mathcal{A}} \alpha_n(x_n)}$.

IV Proposed Upper Bound

The upper bound of the mutual information can be expressed as [6]

$$\bar{I}(X; Y) = \bar{H}(\theta|Y, X) + \bar{H}(Y) - H(Y|X, \theta) - \bar{H}(\theta). \quad (19)$$

In this section, we compute $\bar{H}(\theta|Y, X)$ and $\bar{H}(Y)$ with the framework of circular random variables.

A Compute $\bar{H}(\theta|Y, X)$

This upper bound is computed using the chain property of the information measure and the asymptotic equipartition property as [6]

$$\bar{H}(\theta|Y, X) \approx \frac{1}{M} \sum_{n=1}^M \log_2 \frac{1}{q(\theta_n | x_1^n, y_1^n, \theta_{n+1})}, \quad (20)$$

where $q(\cdot)$ is a distribution that approximates $p(\cdot)$, which is computed as

$$p(e^{j\theta_n} | y_1^n, x_1^n, e^{j\theta_{n+1}}) = \lambda_n p(e^{j\theta_n} | y_1^n, x_1^n) p(e^{j\theta_{n+1}} | e^{j\theta_n}), \quad (21)$$

where λ_n is a constant factor such that the result is a probability density function. The details of computing $p(e^{j\theta_n}|y_1^n, x_1^n)$ is deferred to the end of this subsection. It suffices to state here that we approximate it as $q(e^{j\theta_n}|y_1^n, x_1^n) \approx \mathcal{CN}(e^{j\theta_n}; M_n, V_n)$, and evaluate (21) as

$$p(e^{j\theta_n}|y_1^n, x_1^n, e^{j\theta_{n+1}}) \approx \mathcal{CN}(e^{j\theta_n}; \hat{M}_n, \hat{V}_n), \quad (22)$$

where the mean is

$$\hat{M}_n = \frac{(M_n)(V_n)^{-1} + (e^{j\theta_{n+1}})(\sigma_\theta^2)^{-1}}{(V_n)^{-1} + (\sigma_\theta^2)^{-1}}, \quad (23)$$

and the variance is

$$\hat{V}_n = \frac{1}{(V_n)^{-1} + (\sigma_\theta^2)^{-1}}. \quad (24)$$

Hence, we compute the upper bound of $H(\theta|X, Y)$ using (22) in (20).

The details of computing $q(e^{j\theta_n}|y_1^n, x_1^n)$ is presented next. The distribution $p(e^{j\theta_n}|x_1^n, y_1^n)$ is formulated as

$$p(e^{j\theta_n}|y_1^n, x_1^n) \propto p(e^{j\theta_n}|y_1^{n-1}, x_1^{n-1})p(y_n|x_n, e^{j\theta_n})p(x_n), \quad (25)$$

where $p(e^{j\theta_n}|y_1^{n-1}, x_1^{n-1})$ is formulated as

$$\begin{aligned} p(e^{j\theta_n}|y_1^{n-1}, x_1^{n-1}) &= \int p(e^{j\theta_{n-1}}|x_1^{n-1}, y_1^{n-1}) \\ &\quad \times p(e^{j\theta_n}|e^{j\theta_{n-1}}) de^{j\theta_{n-1}}. \end{aligned} \quad (26)$$

Hence

$$q(e^{j\theta_n}|y_1^n, x_1^n) \propto \mathcal{CN}(e^{j\theta_n}; \bar{M}_n, \bar{V}_n) \mathcal{CN}(y_n; x_n e^{j\theta_n}, \sigma^2) p(x_n), \quad (27)$$

where $q(e^{j\theta_n}|y_1^{n-1}, x_1^{n-1}) \approx \mathcal{CN}(e^{j\theta_n}; \bar{M}_n, \bar{V}_n)$. Further, we use the properties of Gaussian functions [14] to write $q(e^{j\theta_n}|y_1^n, x_1^n) \approx \mathcal{CN}(e^{j\theta_n}; M_n, V_n)$, where the mean is

$$M_n = \frac{(\bar{M}_n)(\bar{V}_n)^{-1} + (\frac{y_n}{x_n})(\frac{\sigma^2}{|x_n|^2})^{-1}}{(\bar{V}_n)^{-1} + (\frac{\sigma^2}{|x_n|^2})^{-1}}, \quad (28)$$

and the variance is

$$V_n = \frac{1}{(\bar{V}_n)^{-1} + (\frac{\sigma^2}{|x_n|^2})^{-1}}. \quad (29)$$

Further, we can approximate (26) as $q(e^{j\theta_n}|y_1^{n-1}, x_1^{n-1}) \approx \mathcal{CN}(e^{j\theta_n}; \bar{M}_n, \bar{V}_n)$, where $\bar{M}_n = M_{n-1}$ and $\bar{V}_n = V_{n-1} + \sigma_\theta^2$.

B Compute $q(Y)$

Here, we compute $q(Y)$ based on the distribution [10] which we write with the framework of circular variables as

$$p(e^{j\theta_n}|y_1^n) = \frac{p(e^{j\theta_n}|y_1^{n-1})p(y_n|e^{j\theta_n})}{\int p(e^{j\theta_n}|y_1^{n-1})p(y_n|e^{j\theta_n}) de^{j\theta_n}}, \quad (30)$$

where the denominator is evaluated as $q(y_n|y_1^{n-1})$. Further,

$$p(e^{j\theta_n}|y_1^{n-1}) = \int p(e^{j\theta_{n-1}}|y_1^{n-1})p(e^{j\theta_n}|e^{j\theta_{n-1}}) de^{j\theta_{n-1}}, \quad (31)$$

and $p(y_n|e^{j\theta_n}) = \sum_{x_n \in \mathcal{A}} p(x_n)p(y_n|x_n, e^{j\theta_n})$. We approximate $p(e^{j\theta_{n-1}}|y_1^{n-1})$ in (31) as $\mathcal{CN}(e^{j\theta_{n-1}}; \dot{M}_{n-1}, \dot{V}_{n-1})$, and then compute $p(e^{j\theta_n}|y_1^{n-1})$ as

$$q(e^{j\theta_n}|y_1^{n-1}) \approx \int \mathcal{CN}(e^{j\theta_{n-1}}; \dot{M}_{n-1}, \dot{V}_{n-1}) \times \mathcal{CN}(e^{j\theta_n}; e^{j\theta_{n-1}}, \sigma_\theta^2) de^{j\theta_{n-1}} \quad (32)$$

$$\approx \mathcal{CN}(e^{j\theta_n}; M_n, V_n), \quad (33)$$

where $M_n = \dot{M}_{n-1}$ and $V_n = \dot{V}_{n-1} + \sigma_\theta^2$. The distribution in (30) is computed as

$$q(e^{j\theta_n}|y_1^n) \propto \mathcal{CN}(e^{j\theta_n}; M_n, V_n) \sum_{x_n \in \mathcal{A}} p(x_n)p(y_n|x_n, e^{j\theta_n}) \quad (34)$$

$$\stackrel{(a)}{\propto} \sum_{x_n \in \mathcal{A}} \alpha_n(x_n) \mathcal{CN}(e^{j\theta_n}; \dot{M}_n(x_n), \dot{V}_n(x_n)), \quad (35)$$

where (a) follows from the properties of Gaussian distributions [14], and

$$\alpha_n(x_n) = p(x_n) \mathcal{CN}(y_n; x_n M_n, \sigma^2 + |x_n|^2 V_n). \quad (36)$$

Further, $\dot{M}_n(x_n)$ describes the mean which is calculated as

$$\dot{M}_n(x_n) = \frac{\left(\frac{y_n}{x_n}\right) \left(\frac{\sigma^2}{|x_n|^2}\right)^{-1} + (M_n)(V_n)^{-1}}{\left(\frac{\sigma^2}{|x_n|^2}\right)^{-1} + (V_n)^{-1}}, \quad (37)$$

and $\dot{V}_n(x_n)$ describes the variance computed as

$$\dot{V}_n(x_n) = \frac{1}{\left(\frac{\sigma^2}{|x_n|^2}\right)^{-1} + (V_n)^{-1}}. \quad (38)$$

The distribution shown in (35) is composed of multiple Gaussian functions and we use the clustering method [15] and approximate it as $q(e^{j\theta_n}|y_1^n) \approx \mathcal{CN}(e^{j\theta_n}; \dot{M}_n, \dot{V}_n)$ with mean $\dot{M}_n = \frac{1}{\beta_n} \sum_{x_n \in \mathcal{A}} \alpha_n(x_n) \dot{M}_n(x_n)$ and variance $\dot{V}_n = \frac{1}{\beta_n} \sum_{x_n \in \mathcal{A}} \alpha_n(x_n) (\dot{V}_n(x_n) + (\dot{M}_n(x_n) - \dot{M}_n)^2)$. Further, $\beta_n = \sum_{x_n \in \mathcal{A}} \alpha_n(x_n)$ is the normalising factor. Therefore, we compute

upper bound as $\overline{H}(y_1^M) \approx \frac{1}{M} \sum_{n=1}^M \log_2 \frac{1}{\sum_{x_n \in \mathcal{A}} \alpha_n(x_n)}$.

C Compute $H(Y|X, \theta)$ and $\overline{H}(\theta)$

The $H(Y|X, \theta)$ is due to the Gaussian noise and hence is computed as $\log_2(\pi e \sigma^2)$. Similarly, we compute $\overline{H}(\theta)$ as $\log_2(\pi e \sigma_\theta^2)$.

V Results

The performance of the proposed methods is obtained with simulations. The information bounds are computed at each signal-to-noise ratio (SNR) using sequences of size 10^5 . The transmitted sequences are selected from uniformly distributed Quadrature Amplitude Modulated (QAM) sets. The Wiener phase noise varies as $\sigma_\theta = 0.125$. For comparison, we use the discrete phase method [5], the Tikhonov-Fourier method [10], and the Kalman method [9]. Furthermore, we include the mutual information of the additive white Gaussian noise channel for constrained and Gaussian inputs. The constrained input refers to the mutual information obtained with discrete input sequences of the modulation format, whereas the Gaussian input refers to sequences which have Gaussian distribution.

Figure 1 shows the performance of the proposed methods for 4-QAM. The proposed parametric lower and upper bounds are denoted as ‘Proposed LB’ and ‘Proposed UB’ accordingly. The proposed lower and upper bounds show similar results as the discrete phase method, which serves as an exact information bound of the channel. The proposed lower bound also shows similar result as the method based on the Kalman filter. Furthermore, the proposed upper bound is shown to be accurate compared to the discrete phase method and the Tikhonov-Fourier method. Figure 2 shows the performance of the proposed methods for 16-QAM. The lower and upper bounds of the proposed methods are shown to be accurate at low-to-moderate SNR compared to the discrete phase method. The accuracy of the upper bound at high SNRs is slightly lower compared to the accuracy at low-to-moderate values of SNR. The upper bound has been approximated based on multiple Gaussian functions which are clustered into one distribution. This clustering of distributions causes the upper bound to be slightly higher than the exact bounds. This would improve with clustering the multiple Gaussian functions in two or more distributions.

In summary, the proposed parameter based methods have shown accurate information bounds at low-to-moderate SNR, whereas very close information bounds have been obtained at high SNR. However, for codes which are commonly used in practice, the proposed algorithms have shown accurate bounds. Since the information bounds only require the mean and variance of the Gaussian function, the proposed methods have significantly lower complexity than the discrete based method. The computation of the information rates for higher modulation formats, such as the 32-QAM and 64-QAM, follows a similar approach as the one presented in this work. The discrete method would require more

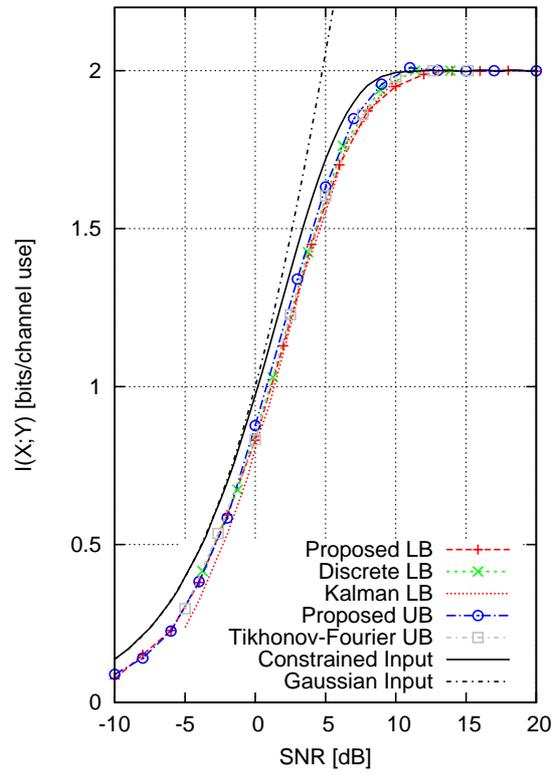


Figure 1: Information bounds of the proposed methods 4-QAM.

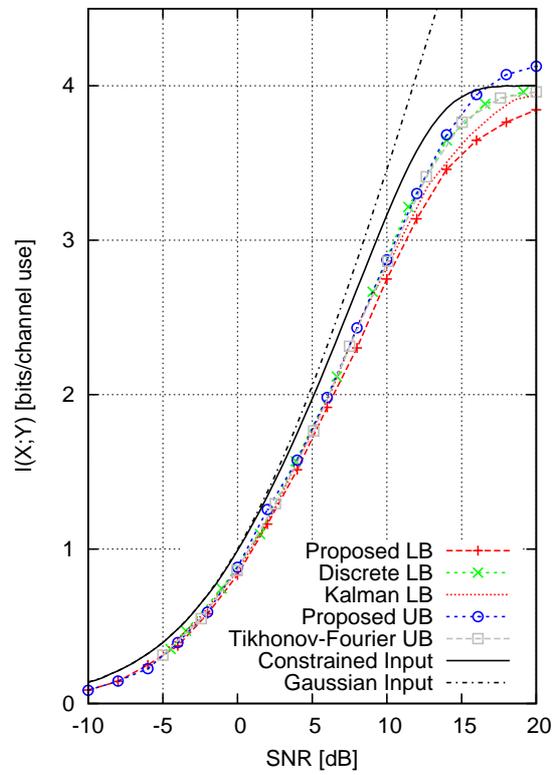


Figure 2: Information bounds of the proposed methods 16-QAM.

sampled values than the 16-QAM. The proposed lower bound would compute the information rates at much lower complexity than the discrete method. The upper bound would require two or more clusters of Gaussian functions to achieve good accuracy. In general, the complexity of this process is much lower than the discrete method. This makes the proposed methods an attractive choice to characterise the performance of communication systems based on Wiener phase noise channels.

VI Conclusion

In this paper, we have proposed parameter based methods to compute the information bounds of the Wiener phase noise channel. The methods represent the phase noise as a circular random variable and compute the information bounds based on parameters of the Gaussian distribution. Accurate results have been obtained at low-to-moderate signal-to-noise ratios, whereas very close information bounds have been shown at high signal-to-noise ratios.

References

- [1] R.-J. Essiambre, G. Kramer, P. J. Winzer, G. J. Foschini, and B. Goebel, “Capacity limits of optical fiber networks,” *IEEE Journal of Lightwave Technology*, vol. 28, no. 4, pp. 662–701, 2010.
- [2] G. J. Foschini and G. Vannucci, “Characterizing filtered light waves corrupted by phase noise,” *IEEE Transactions on Information Theory*, vol. 34, no. 6, pp. 1437–1448, 1988.
- [3] M. Magarini, A. Spalvieri, F. Vacondio, M. Bertolini, M. Pepe, and G. Gavioli, “Empirical modeling and simulation of phase noise in long-haul coherent optical transmission systems,” *Optics Express*, vol. 19, no. 23, pp. 22 455–22 461, 2011.
- [4] D.-M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, “Simulation-based computation of information rates for channels with memory,” *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3498–3508, 2006.
- [5] L. Barletta, M. Magarini, and A. Spalvieri, “Estimate of information rates of discrete-time first-order Markov phase noise channels,” *IEEE Photonics Technology Letters*, vol. 23, no. 21, pp. 1582–1584, 2011.
- [6] —, “The information rate transferred through the discrete-time Wiener’s phase noise channel,” *IEEE Journal of Lightwave Technology*, vol. 30, no. 10, pp. 1480–1486, 2012.
- [7] M. Martalò, C. Tripodi, and R. Raheli, “On the information rate of phase noise-limited communications,” in *ITA*. Citeseer, 2013, pp. 1–7.
- [8] J. Dauwels and H.-A. Loeliger, “Computation of information rates by particle methods,” in *IEEE International Symposium on Information Theory*, 2004, p. 178.
- [9] L. Barletta, M. Magarini, and A. Spalvieri, “A new lower bound below the information rate of Wiener phase noise channel based on kalman carrier recovery,” *Optics Express*, vol. 20, no. 23, pp. 25 471–25 477, 2012.
- [10] L. Barletta, “Computing the information rate of discrete-time Wiener phase noise channels by parametric bayesian tracking,” *Optics Express*, vol. 24, no. 23, pp. 26 612–26 617, 2016.
- [11] M. P. Yankov, T. Fehenberger, L. Barletta, and N. Hanik, “Low-complexity tracking of laser and nonlinear phase noise in WDM optical fiber systems,” *IEEE Journal of Lightwave Technology*, vol. 33, no. 23, pp. 4975–4984, 2015.
- [12] S. H. Rezenom and F. Takawira, “Iterative low complexity algorithm for LDPC systems in the presence of phase noise,” *IEEE Wireless Communication Letters*, vol. 6, no. 6, pp. 794–797, 2017.

- [13] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [14] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann, 1988.
- [15] J. Goldberger and S. T. Roweis, “Hierarchical clustering of a mixture model,” in *Advances in Neural Information Processing Systems*, 2005, pp. 505–512.

III

Conclusion

Conclusion and Future work

1 Conclusion

In this thesis, various graphical algorithms have been proposed to detect signals in the presence of device uncertainties such as carrier frequency offset and Wiener phase noise.

In Paper 1, the effect of message scheduling has been studied in the detection of OFDM signals affected by carrier frequency offset and Wiener phase noise. The proposed scheduling algorithm has been shown to converge faster than the existing scheduling methods.

In Paper 2, the circular characteristics of phase noise has been exploited to propose an algorithm that computes the belief propagation algorithm using approximate closed-form expressions based on parametric message representation. The proposed algorithm has been used to detect LDPC signals in the presence of Wiener phase noise. The results show that the proposed algorithm achieves similar performance as the existing algorithm with the lowest complexity. Further, the proposed algorithm has improved the complexity of the existing algorithm.

In Paper 3, an algorithm based on the circular random variables has been proposed to detect OFDM signals in the presence of frequency offset and Wiener phase noise. The proposed algorithm has shown bit-error rate performance close to the ideal algorithm. Further, the complexity the proposed algorithm has been shown to be significantly lower than the existing algorithm.

In Paper 4, algorithms based on minimising the free-energy function have been proposed to detect LDPC codes in the presence of Wiener phase noise. The free energy function of the proposed algorithms have been formulated using the combined BP and URWBP algorithm, and the combined MF and URWBP algorithm. The finding highlights the importance of combining two or more graphical algorithms to outperform the performance of existing algorithms.

In Paper 5, efficient low complexity methods have been proposed to compute the lower and upper mutual information bounds of the Wiener phase noise channel. The proposed methods have exploited the phase noise distributions to be expressed with the circular random variables. The lower and upper mutual information bounds have been obtained using parameters of the Gaussian density functions. The results have shown similar performance as the existing high complexity methods at a significantly low implementation complexity.

2 Future work

The proposed algorithm using circular random variables has very low complexity, and this advantage over existing algorithms has a direct impact on the Digital Video Broadcasting (DVB) standards for satellite communications [1]–[3]. The proposed algorithm can be extended to use the system setup described in these standards such as higher modulation formats and longer length of the LDPC codes.

The algorithms based on the free energy function and the circular random variables can also be extended to other problems in communication systems such as the cooperative wireless localisation problem [4]. The algorithm can model the wireless nodes depending on the degree of connectivity to neighbouring nodes. The combined BP and URWBP algorithm can be used to solve this problem where the BP would be used for low-degree nodes and the URWBP can be used for high-degree nodes. Another aspect could model the position information using the circular random variables instead of existing approaches. The algorithms that minimise the energy functions could also be used for this purpose. The aim will be to establish whether these approaches improve the complexity of existing algorithms [4]–[7].

The applications envisioned for next generation wireless communication systems require very short latency, better spectral efficiency, and extremely high data rates compared with the existing systems [8]. The connection of many interconnected devices, and their potential applications, creates a need for new functionality in communication systems [9]. These requirements could not be met using the existing technologies in their present form. Therefore, these factors are motivating many researchers to experiment with new modulation methods that increase the data rates significantly, lower the latency, support many interconnected devices, and improve spectral efficiency [10], [11].

There have been various approaches to achieve these requirements. Some of these approaches improve upon the existing communication systems based on the OFDM system, which has disadvantages such as longer symbol duration, strong out-of-band frequency spectrum and low spectral efficiency due to the cyclic prefix. These disadvantages have been counteracted using techniques such as the Filter Bank Multi-Carrier (FBMC) [12], Generalised Frequency Division Multiplexing (GFDM) [13], [14], Universal Filter Multi-Carrier (UFMC) [15], etc. Each of these techniques has been shown to improve the weaknesses of existing OFDM systems and are considered to be candidates for future generation wireless communication systems.

To satisfy the requirements of significantly high data rates, there have been proposals to use the Extremely High Frequency (EHF) band, also referred to as the mm-Wave band [16]. The use of this frequency band for communication systems has gained research attention and modulation techniques are being developed. Further, the use of large number of antennas at the base station has been proposed to increase the available data rates [17]. There is an increasing research interest on non-orthogonal multiple access techniques where the radio resource is shared among many users [18]–[20].

These advances bring new type of system models which are still affected by carrier frequency offset and phase noise. In fact, the high frequency band makes the effect of frequency offset and phase noise more significant compared to the existing systems, and there are very few studies that have proposed algorithms for these systems. Hence, the graphical algorithms developed in this thesis could be modified to the system model of future communications systems.

References

- [1] E. Casini, R. D. Gaudenzi, and A. Ginesi, “DVB-S2 modem algorithms design and performance over typical satellite channels,” *International Journal of Satellite Communications and Networking*, vol. 22, no. 3, pp. 281–318, 2004.
- [2] “ETSI EN 302 307-1 V1.4.1: Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications; part I: DVB-S2,” 2014.
- [3] “ETSI EN 302 307-2 V1.4.1: Digital Video Broadcasting (DVB); Second generation framing structure, channel coding and modulation systems for broadcasting, interactive services, news gathering and other broadband satellite applications; part II: DVB-S2X,” 2015.
- [4] H. Wymeersch, J. Lien, and M. Z. Win, “Cooperative localization in wireless networks,” *Proceedings of the IEEE*, vol. 97, no. 2, pp. 427–450, 2009.
- [5] J. Shen, A. F. Molisch, and J. Salmi, “Accurate passive location estimation using TOA measurements,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 6, pp. 2182–2192, 2012.
- [6] W. Yuan, N. Wu, Q. Guo, X. Huang, Y. Li, and L. Hanzo, “TOA-based passive localization constructed over factor graphs: A unified framework,” *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 6952–6965, 2019.
- [7] W. Yuan, N. Wu, B. Etxzlinger, Y. Li, C. Yan, and L. Hanzo, “Expectation-maximization-based passive localization relying on asynchronous receivers: Centralized versus distributed implementations,” *IEEE Transactions on Communications*, vol. 67, no. 1, pp. 668–681, 2018.
- [8] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, “What will 5G be?” *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1065–1082, 2014.
- [9] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, “A survey of millimeter wave communications (mmWave) for 5G: opportunities and challenges,” *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.
- [10] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, “Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency,” *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, 2014.

- [11] S. Buzzi, C. D’Andrea, T. Foggi, A. Ugolini, and G. Colavolpe, “Spectral efficiency of MIMO millimeter-wave links with single-carrier modulation for 5G networks,” in *Proceedings of the 20th International ITG Workshop on Smart Antennas (WSA)*, 2016, pp. 1–8.
- [12] B. Farhang-Boroujeny, “Filter bank spectrum sensing for cognitive radios,” *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1801–1811, 2008.
- [13] G. Fettweis, M. Krondorf, and S. Bittner, “GFDM-generalized frequency division multiplexing,” in *IEEE Vehicular Technology Conference*, 2009, pp. 1–4.
- [14] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, “Generalized frequency division multiplexing for 5th generation cellular networks,” *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3045–3061, 2014.
- [15] V. Vakilian, T. Wild, F. Schaich, S. ten Brink, and J.-F. Frigon, “Universal-filtered multi-carrier technique for wireless systems beyond lte,” in *IEEE Globecom Workshop*, 2013, pp. 223–228.
- [16] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, “Millimeter wave mobile communications for 5G cellular: It will work!” *IEEE Access*, vol. 1, pp. 335–349, 2013.
- [17] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive MIMO for next generation wireless systems,” *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [18] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, “Non-orthogonal multiple access (NOMA) for cellular future radio access,” in *2013 IEEE 77th vehicular technology conference (VTC Spring)*, 2013, pp. 1–5.
- [19] M. Vaezi, Z. Ding, and H. V. Poor, *Multiple access techniques for 5G wireless networks and beyond*. Springer, 2019.
- [20] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-Lin, and Z. Wang, “Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends,” *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, 2015.