Thesis for the degree of Doctor of Philosophy

On the Impact of Phase Noise in Communication Systems - Performance Analysis and Algorithms

RAJET KRISHNAN

CHALMERS

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology

Gothenburg 2015

Krishnan, Rajet
On the Impact of Phase Noise in Communication Systems - Performance Analysis and Algorithms.

Department of Signals and Systems
Technical Report No. 3849
ISBN: 978-91-7597-168-1

Communication Systems Group
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Gothenburg, Sweden
Telephone: + 46 (0)31-772 1000

This thesis has been prepared using LaTeX.

Printed by Chalmers Reproservice,
Gothenburg, Sweden, April 2015.

To science,
to those who have dedicated their lives to the pursuit of knowledge and the truth,
and to those who have sacrificed their lives for equality, freedom and justice.

Abstract

The mobile industry is preparing to scale up the network capacity by a factor of 1000x in order to cope with the staggering growth in mobile traffic. As a consequence, there is a tremendous pressure on the network infrastructure, where more cost-effective, flexible, high speed connectivity solutions are being sought for. In this regard, massive multiple-input multiple-output (MIMO) systems, and millimeter-wave communication systems are new physical layer technologies, which promise to facilitate the 1000 fold increase in network capacity. However, these technologies are extremely prone to hardware impairments like phase noise caused by noisy oscillators. Furthermore, wireless backhaul networks are an effective solution to transport data by using high-order signal constellations, which are also susceptible to phase noise impairments.

Analyzing the performance of wireless communication systems impaired by oscillator phase noise, and designing systems to operate efficiently in strong phase noise conditions are critical problems in communication theory. The criticality of these problems is accentuated with the growing interest in the new physical layer technologies, and the deployment of wireless backhaul networks. This forms the main motivation for this thesis where we analyze the impact of phase noise on the system performance, and we also design algorithms in order to mitigate phase noise and its effects.

First, we address the problem of maximum a posteriori (MAP) detection of data in the presence of strong phase noise in single-antenna systems. This is achieved by designing a low-complexity joint phase-estimator data-detector. We show that the proposed method outperforms existing detectors, especially when high order signal constellations are used. Then, in order to further improve system performance, we consider the problem of optimizing signal constellations for transmission over channels impaired by phase noise. Specifically, we design signal constellations such that the error rate performance of the system is minimized, and the information rate of the system is maximized. We observe that these optimized constellations significantly improve the system performance, when compared to conventional constellations, and those proposed in the literature.

Next, we derive the MAP symbol detector for a MIMO system where each antenna at the transceiver has its own oscillator. We propose three suboptimal, low-complexity algorithms for approximately implementing the MAP symbol detector, which involve joint phase noise estimation and data detection. We observe that the proposed techniques significantly outperform the other algorithms in prior works. Finally, we study the impact of phase noise on the performance of a massive MIMO system, where we analyze both uplink and downlink performances. Based on rigorous analyses of the achievable rates, we provide interesting insights for the following question: how should oscillators be connected to the antennas at a base station, which employs a large number of antennas?

Keywords: Oscillator, phase noise, maximum likelihood (ML) detection, maximum a posteriori (MAP) detection, extended Kalman filter (EKF), constellations, symbol error probability, mutual information, sum-product algorithm (SPA), factor graph, variational Bayesian method, multiple-input multiple-output (MIMO), massive MIMO, random matrix theory, free probability.

List of Publications

Appended papers

This thesis is based on the following papers:

[A] Rajet Krishnan, M. Reza Khanzadi, T. Eriksson, T. Svensson, "Soft metrics and their Performance Analysis for Optimal Data Detection in the Presence of Strong Oscillator Phase Noise," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2385-2395, Jun. 2013.

[B] Rajet Krishnan, A. Graell i Amat, T. Eriksson, G. Colavolpe, "Constellation Optimization in the Presence of Strong Phase Noise," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5056-5066, Dec. 2013.

[C] Rajet Krishnan, G. Colavolpe, A. Graell i Amat, T. Eriksson, "Algorithms for Joint Phase Estimation and Decoding for MIMO Systems in the Presence of Phase Noise," accepted for publication in *IEEE Trans. Signal Process.* , Jan. 2015.

[D] Rajet Krishnan, M. R. Khanzadi, N. Krishnan, A. Graell i Amat, T. Eriksson, Yongpeng Wu, Robert Schober, "Linear Precoding in the Presence of Phase Noise in Massive MIMO Systems: A large scale analysis," under review, *IEEE Trans. Veh. Tech.*, Jan. 2015.

[E] Rajet Krishnan, M. R. Khanzadi, N. Krishnan, N. Mazzali, A. Graell i Amat, T. Eriksson, G. Colavolpe, "On the Impact of Oscillator Phase Noise in Uplink Transmission for Massive MIMO-OFDM Systems," to be resubmitted, *IEEE Signal Process. Lett.*, Apr. 2015.

Ongoing works and other papers

The following papers are not appended to this thesis, either due to content overlapping with the appended papers, or due to content which is not related to this thesis.

[a] M. R. Khanzadi, Rajet Krishnan, *et al.* "Uplink and downlink performance of massive MIMO systems in the presence of phase noise, reduced RF chains, and finite-resolution signal converters".

[b] A. Papazafeiropoulos, Rajet Krishnan, *et al.* "Impact of channel aging due to phase noise and user mobility on the performance of massive MIMO systems".

[c] Karl Rundstedt, Rajet Krishnan, *et al.* "Measurements, modeling and performance analysis of line-of-sight MIMO microwave channels".

[d] Arni Alfredsson, Rajet Krishnan, *et al.* "Algorithms for phase noise compensation and data detection for optical communication systems in the presence of optical phase noise".

[e] M. Reza. Khanzadi, Rajet Krishnan, T. Eriksson, "Estimation of Phase Noise in Oscillators with Colored Noise Sources," *IEEE Commun. Lett.*, vol. 17, no. 11, pp. 2160-2163, Nov. 2013.

[f] M. Reza. Khanzadi, Rajet Krishnan, T. Eriksson, "On the Capacity of the Wiener Phase Noise Channel: Bounds and Capacity Achieving Distributions," to be submitted, *IEEE Trans. Commun.*, Apr. 2015.

[g] Rajet Krishnan, Hani Mehrpouyan, Thomas Eriksson, and Tommy Svensson , "Optimal and Approximate Methods for Detection of Uncoded Data with Carrier Phase Noise," In *Proc. Global Commun. Conf. (GLOBECOM 2011)*, Houston, pp. 1-6, 5-9 Dec. 2011.

[h] Rajet Krishnan, M. R. Khanzadi, L. Svensson, Thomas Eriksson, and Tommy Svensson, "Variational Bayesian Framework for Receiver Design in the Presence of Phase Noise in MIMO Systems," In *Proc. IEEE Wireless Commun. Net. Conf. (WCNC 2012)*, Paris, France, pp. 34-39, 1-4 Apr. 2012.

[i] M. R. Khanzadi, Rajet Krishnan, and Thomas Eriksson, "Effect of Synchronizing Coordinated Base Stations on Phase Noise Estimation," In *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Vancouver, Canada, pp. 4938-4942, 26-31 May 2013.

[j] M. R. Khanzadi, Rajet Krishnan, Dan Kuylenstiern and Thomas Eriksson, "Oscillator Phase Noise and Small-Scale Channel Fading in Higher Frequency Bands," *Globecom 2014 Workshop - Mobile Communications in Higher Frequency Bands, 2014*, accepted for publication.

[k] M. R. Khanzadi, Rajet Krishnan, and Thomas Eriksson, "Differential signaling for multiple-antenna phase noise channels with common and separate oscillator configurations," to be submitted to *Globecom 2015*.

[l] Rajet Krishnan, M. R. Khanzadi, N. Krishnan, A. Graell i Amat, T. Eriksson, Yongpeng Wu, Robert Schober, "A large-scale performance analysis of linear precoding in the presence of phase noise in massive MIMO systems," accepted, *ICC 2015*, Mar. 2015.

[m] Rajet Krishnan, G. Colavolpe, A. Graell i Amat, T. Eriksson, "Sum-product algorithms for joint phase estimation and decoding for MIMO systems in the presence of phase noise," to be submitted to *Globecom 2015*.

Acknowledgements

As my journey of life to destination unknown continues, I wish to pause for a moment to graciously thank all those who have made my experience in the last four and a half earth years a great one. Thomas, I don't think that I would be able to express in words as to what you mean to me. Thanks for everything—for accepting me as your PhD student, inspiring me to do my best, understanding me, my dreams and my ambitions, supporting me—absolutely everything! What makes working with you memorable (along with our brainstorming sessions) is that I could blurt out the most stupid things, commit the most frivolous mistakes, and you still would be non-judgmental and forgiving. Even going forward, you know that I will come back to you for advice whenever I hit the wall. So, you are not done with me yet!

Tommy, thank you for recruiting me, convincing me to come to Chalmers and work with Thomas. I have hugely benefitted from the tips and the advices that you have given me over these years. I have thoroughly enjoyed our brainstorming sessions on research, technology, entrepreneurship and life! I really appreciate the big picture imagery that you offer when I get stuck in puny problems. Alex, I would like to thank you for the countless hours that you spent in reading my papers, and helping me hone my writing skills. I am really grateful for all the guidance that you have offered, and the constant support that you have provided in the last two and a half years. I have always loved it when I rant about my aspirations and ambitions, and thanks for not telling me to stop—those rants helped me a lot to understand myself! Lennart, I cannot thank you enough for all the discussions that we have had on the fascinating topics in Bayesian statistics and graphical inference. Most importantly, hats off to your patient hearing to my bombastic ideas. I have learnt a lot from you, and though many of our interesting ideas and discussions have remained on the black-board, I hope to revisit them in the future.

Special thanks to Erik for his leadership and guidance to the Commsys group – you have always gone the extra mile to offer us a terrific work environment where we could work and collaborate effectively, and be productive and creative. Special thanks to Agneta, Malin, Lars, Natasha, and Karin – working at Commsys would not have been so good if not for all of you. I would like to thank Robert Schober at FAU, Erlangen, and Giulio Colavolpe at the University of Parma, for hosting me and collaborating with me – thanks for your friendship, the rich and intense research discussions, and a big toast to the challenging problems that we worked on. It was great working with you! Thanks to my collaborators at Ericsson for giving me the opportunity to work and collaborate on the various challenging technical problems in the industry.

Now over to my research buddy, and 'partner in crime', the one and only Reza! We have done so much over the last 54 months; some of which can be quantified as the papers that we wrote together, and most of which would sediment as lifetime-memories and wonderful experiences. Thanks for all those random, serious, and silly discussions that we routinely conducted with so much fervor, our passionate visions of our future, and much, much more. I have always admired your drive for innovation backed by your diverse technical skill-set, which includes handling mathematical problems to building real stuffs. To all my wonderful friends at the department—both current and former members—thanks for your friendship. I have always looked forward to coming to office everyday, interacting with all of you, and having good times together. Toast to those good times, which will remain close to my heart going forward. I have been always inspired by you, and I have learnt a lot from you.

To the most important person of my life, my wife and soulmate, Swathi – I just love the way you support me when I am down in my spirits; I just love the way you lend me your ears patiently when I jibber incessantly about work and my dreams everyday; I just love the way you make me feel; I just love the way you love me! Thanks for the good times, for accompanying me in this journey, and for being always there for me through thick and thin. From my heart, I love you. I wish and hope that all your dreams and wishes come true, and I will always be there for you. Thanks to my parents and siblings for all their love, care, and laying the foundation for what I am, and where I am today, and the values that I stand for – you have always kept me going. Thanks to my in-laws, extended family and my great friends who have loved me, believed in me and supported me. As Swathi and I selfishly propagate our genes, and in consequence, as we wait eagerly for our baby to enter this beautiful world and our beautiful lives, I wish to leave a small note – Mama and Papa are waiting for you with all our love to greet you, see you, hear you, feel you, and take you along in our journey. May all your dreams and wishes come true.

So, as my journey continues on the untrodden paths; over uncharted mountains; in unexplored oceans; on the harsh wintry snow-lands, and the blistering-hot sand-dunes, I will remember you all, the good times that we had, and the inspiration that you have given me...

Rajet Krishnan
Gothenburg, April 2015

Contents

II Included papers 41

Part I

Overview

Chapter 1

Introduction

Since the landmark paper by Shannon [1], substantial research has been done in order to design communications systems that operate close to the ultimate performance limit, i.e., the channel capacity, with an arbitrary small probability of error. Particularly in single-input single-output (SISO) and multiple-input multiple-output (MIMO) wireless systems, much of these efforts have been based on several idealized assumptions like perfect channel state information (CSI), perfect synchronization, ideal hardware, untethered implementation complexity, and much more.

As a result of the idealized assumptions, there is a significant gap between the performance realized in state-of-the-art communication systems, and that promised by theory. A major contributor to this performance gap are the impairments due to non-ideal hardware components used in communication transceivers [2]. In particular, erroneous CSI is a substantial source of performance loss—about $2-3$ dB loss is incurred even with the best CSI estimates in LTE-A, and upto 14% throughput is lost owing to the use of pilots. Hardware impairments adversely affect the CSI quality, and in particular, errors due to phase noise caused by noisy oscillators used in the transceivers can result in significant performance degradation [3]. Some of the other important hardware impairments that drastically affect the system performance include amplifier nonlinearities, and the quantization noise caused by finite-resolution signal-converters [4].

Oscillators are central to the design of wireless communication systems, and they should be accurate, inexpensive, and compact. They provide the carrier signals required for passband transmission, and also the reference clock signal for purposes like sampling and synchronization. All practical oscillators suffer from phase noise. Thus, when information is conveyed from a source (transmitter) to a destination (receiver), random time-varying phase variations manifest in the signal obtained at the receiver. Random time-varying phase variations result in rotation of the transmitted information. Furthermore, phase noise causes the effective channel phase to drift randomly between the time instant that a pilot symbol is transmitted/received for CSI acquisition, and the instant that a data symbol is transmitted/received. Thus, the actual effective channel phase during data transmission can become significantly different from the CSI acquired. This is detrimental given that many communication systems are designed to operate synchronously and coherently [2]. If the issues pertaining to phase noise are not appropriately addressed, it can result in undesirably high error rates.

1.1 Aim and Flow of the Thesis

This thesis analyzes the impact of random phase noise on the performance of communication systems, and investigates methods for compensating them. Broadly, we try to address the following questions in the papers that are appended to this thesis:

[a] In the presence of oscillator phase noise, how can we systematically derive low-complexity receiver algorithms for SISO and MIMO systems that are (near) optimal in performance? This question is addressed in Papers A and C [5, 6].

[b] How can we optimize signal constellations for transmission over a channel impaired by random phase noise? This problem is investigated in Paper B [7].

[c] What is the impact of phase noise on the achievable rate of a massive MIMO system? How should oscillators be connected to the antennas in a massive MIMO base station? These problems are studied in Papers D and E [8, 9].

In order to comprehensively answer the aforementioned questions, it is imperative to understand the source of phase noise, and its impact on the system performance. Therefore, in the introductory chapters of this thesis, we will review prior works that are relevant to the following questions:

- What is the source of oscillator phase noise, and how is it mathematically modeled?
- What is the optimal receiver structure (i.e., the optimal detector) in the presence of random phase noise, which minimizes the symbol error probability (SEP) performance in a communication system? Furthermore, what are the bounds for estimating random phase noise, which is required in order to minimize the SEP performance of a communication system?
- How can error correcting codes be designed in order to improve system performance when impaired by phase noise?
- What is the capacity of channels impaired by phase noise?

1.1.1 Thesis Outline

The thesis is organized as follows. In Chapter 2, we review the phase noise generation mechanism in the oscillators used in communication transceivers. Then, we discuss the system models that capture the effects of phase noise in SISO, MIMO, and massive MIMO systems in Chapter 3. In Chapter 4, we present an assortment of the mathematical tools and results, which are used in this thesis. Specifically, we discuss about the application of Bayesian inference methods such as the sum-product algorithm (SPA), and the variational Bayesian (VB) method. These methods are used for designing detectors for joint phase noise estimation and data detection in SISO and MIMO systems. Furthermore, we discuss the Shannon capacity of a SISO phase noise channel, and the basics of random matrix theory (RMT), which are employed for analyzing massive MIMO systems. In Chapter 5, we review prior work related to the design of communication systems that are impaired by phase noise. Finally, we summarize our papers, and the main contributions in Chapter 6.

Chapter 2

Phase Noise in Oscillators

An oscillator is an autonomous electronic circuit, which ideally produces a periodic, oscillating electric signal at a precise frequency. This frequency is commonly used to provide clock signals for timing and frequency synchronization, and carrier signals for passband transmission and reception in communication systems. Oscillators used in communication transceivers are imperfect, in that their output signals are affected by random amplitude and phase instabilities. The signal at the output is written as

$$v(t) = (1 + a(t)) \cos(2\pi f_{\text{osc}} t + \theta(t)), \tag{2.1}$$

where f_{osc} is the center frequency of the oscillator, $a(t)$ represents the random amplitude variations, and $\theta(t)$ denotes the random phase variations. In the time domain, the instabilities manifest as a random jitter in the zero-crossings of the desired signal as shown in Fig. 2.1 (a). In the frequency domain, a spectrum of noise around f_{osc} appears as shown in Fig. 2.1 (b). The amplitude perturbations are typically attenuated by a limiting mechanism in the oscillator circuitry, and hence can be ignored [10]. Phase noise in the signal from the oscillators is the focus of this thesis (2.1). In the remainder of this chapter, we will briefly review the various sources and models for phase noise from a noisy oscillator.

2.1 Noise Sources in an Oscillator

The phase of the oscillator signal is affected by a number of noise sources. Broadly speaking, these sources can be categorized as short-term instabilities, deterministic instabilities and long-term instabilities [11]. Short-term instabilities, which typically last for a few seconds, are mainly caused by the following noise sources:

- Thermal Noise - This refers to the memoryless white noise caused by random motion of thermally excited electrons, and its power is equal to kTB, where k is the Boltzman constant, T is the absolute temperature in Kelvin, and B is the 3-dB noise bandwidth [11].
- Colored Noise - This is the spectral noise dominated by low-frequency components that mixes with frequencies close to the center frequency of the oscillator [12]. The instantaneous fluctuations resulting from this noise depends on its past, and therefore has memory. This noise source is primarily dominated by the $1/f$ Noise or flicker noise.

The main deterministic sources of oscillator noise are the following [13]:

- Power supply feed-through and other interfering sources - Coupling can happen between the oscillator signal and the other signals present in an oscillator circuitry. This can result in amplitude/phase modulation of the output signal from the oscillator.

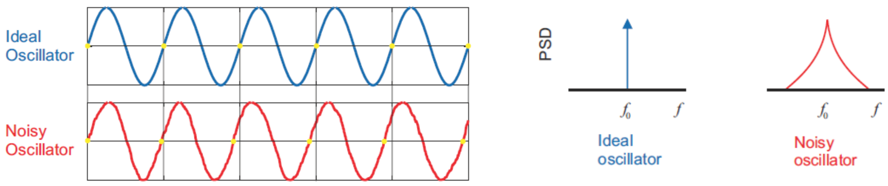


Figure 2.1: Signal output of an oscillator: (a) Time-domain signal, (b) Power spread in the frequency-domain.

- Spurious signals - Generally, an oscillator is designed to have just one feedback path for phase correction, and to generate the desired output signal. However, several feedback paths may exist, which may in turn result in spurious output signals.

Long term instabilities occur in oscillators due to aging of the constituent resonator material. Typically, these are slow variations that occur over hours, days, months, or even years, and are therefore less critical. In the sequel, we present statistical models for the phase noise process for the various noise sources in the oscillator.

2.2 Phase Noise Models

Consider a noisy oscillator operating at f_{osc}, and affected by white noise and colored noise sources as described before. Let $\tilde{\theta}(t)$ represent the sum of all these noise processes. Then, the phase noise at the output signal of the oscillator is given as [14]

$$\theta(t) \propto \int_0^t \tilde{\theta}(t_1)\,\mathrm{d}t_1, \tag{2.2}$$

where $\tilde{\theta}(t_1)$ is a Gaussian process by the central limit theorem. Furthermore, the variance of $\theta(t)$ in (2.2) increases with time [14]. In other words, *the phase noise in an oscillator is an accumulative Gaussian process that results from integrating both the white and colored noise perturbations over time.* The power spectral density (PSD) of the phase noise process $\theta(t)$ in (2.2) is approximately [15]

$$S_\theta(f) \propto \frac{k_2}{f^2} + \frac{k_3}{f^3}, \tag{2.3}$$

where k_2 and k_3 are positive constants that depend on the quality of the oscillator. The PSD of a real oscillator operating at $f_{\text{osc}} = 9.85$ GHz is presented in Fig. 2.2, where $1/f^2$ and $1/f^3$ dependencies of the oscillator PSD are evident.

Now consider the phase noise during a time interval τ, and define

$$\Delta(\tau) \triangleq \theta(t+\tau) - \theta(t) \propto \int_t^{t+\tau} \tilde{\theta}(t_1)\,\mathrm{d}t_1. \tag{2.4}$$

Here $\Delta(\tau)$ refers to the phase noise increment, which is described as the phase noise that has accumulated over the time interval τ. The increment process in (2.4) is also called an *innovation* process. As shown in [17], the increment process is stationary and Gaussian, and its variance is given as

$$\sigma_\Delta^2(\tau) = \int_{-\infty}^{\infty} S_\theta(f) 4 \sin(\pi f \tau)^2 \mathrm{d}f. \tag{2.5}$$

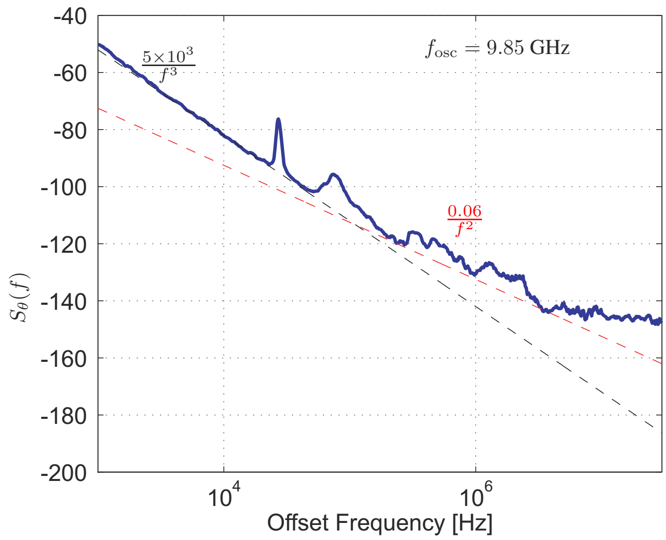


Figure 2.2: PSD of a real oscillator from measurements [16].

When the cumulative noise process $\theta(t)$ in the oscillator is assumed to be only *white* and Gaussian, then the phase noise $\theta(t)$ defined in (2.2) is a *continuous Wiener process* [14]. Furthermore, the variance of the increment process in (2.4) reduces to

$$\sigma_{\Delta}^2(\tau) = 4\pi^2 K_{\mathrm{w}} \tau, \tag{2.6}$$

where K_{w} is a constant, which depends on the cumulative white noise processes in the oscillator. For the remainder of the thesis, we will assume that the oscillator has only white noise sources, and $\theta(t)$ is a continuous Wiener process [14]. This assumption is widely used, and is reasonable since in many practical oscillators, white noise sources are dominant when compared to colored sources.

Chapter 3

Communication System Models in the Presence of Phase Noise

In this chapter, we mathematically model the effect of phase noise in a communication system, and, we show the distortion caused by phase noise on the transmitted signal. To this end, first, we describe the effect of phase noise on a SISO system by analyzing the received signal after matched filtering and Nyquist sampling. Then, we briefly discuss the effect of phase noise on the information signal in the presence of channel fading, and additive white Gaussian noise (AWGN). Finally, we review the basics of MIMO and massive MIMO systems [18, 19], and discuss the impact of phase noise on these systems. Within the scope of MIMO systems, we examine different oscillator setups at the transceiver, and describe the effect of phase noise in the considered setup.

3.1 SISO System Model With Phase Noise

Consider a SISO link, and define the information signal $c(t)$ as

$$c(t) = \sum_{l=0}^{L-1} c_l p(t - lT_\mathrm{s}), \tag{3.1}$$

where T_s is the symbol period, $p(\cdot)$ is a bandlimited square root Nyquist pulse [20], and L is the number of information symbols transmitted. The symbols c_k in (3.1) are drawn from the signal constellation $\mathcal{M} = \{c_i\} \, \forall \, i \in \{1, ..., C\}$, where C is the size of the constellation. Using the signal from the oscillator at the transmitter, $c(t)$ is up-converted to obtain the passband information signal as [20]

$$c_\mathrm{pb}(t) = \Re\{\sqrt{2}c(t)\mathrm{e}^{j(2\pi f_\mathrm{osc} t + \phi(t))}\}, \tag{3.2}$$

where $\Re\{\cdot\}$ denotes the real part of a complex number, and $\phi(t)$ is the Wiener phase noise process caused by the oscillator at the transmitter.

The passband signal $c_\mathrm{pb}(t)$ is transmitted from the source to the destination, and is further affected by phase noise and AWGN at the receiver. Specifically, let $r_\mathrm{pb}(t)$ denote the passband signal received at the destination, where

$$r_\mathrm{pb}(t) = c_\mathrm{pb}(t) + \tilde{n}_\mathrm{pb}(t), \tag{3.3}$$

and $\tilde{n}_{\mathrm{pb}}(t)$ is the passband AWGN process with double-sided noise PSD denoted by N_0. The passband signal $r_{\mathrm{pb}}(t)$ is down-converted to base-band at the receiver as

$$r'(t) = \Re\{\sqrt{2}r_{\mathrm{pb}}(t)\mathrm{e}^{j(2\pi f_{\mathrm{osc}}t+\varphi(t))}\}, \tag{3.4}$$

where $\varphi(t)$ is the phase noise caused by the oscillator at the receiver. The signal $r'(t)$ is then low-pass filtered to obtain $r(t)$, which is written as

$$r(t) = c(t)e^{j\theta(t)} + \tilde{n}'(t), \tag{3.5}$$

where $\theta(t) = \phi(t) + \varphi(t)$, and $\tilde{n}'(t)$ is an AWGN process with double-sided noise PSD N_0, which corresponds to the complex envelope of $\tilde{n}_{\mathrm{pb}}(t)$. The noise processes $\theta(t)$, $\tilde{n}'(t)$ are independent of each other, and also of $c(t)$.

The received signal (3.5) is passed through a matched filter $p^*(-t)$ and sampled at the Nyquist rate, T_{s}, as

$$\begin{aligned} r(kT_{\mathrm{s}}) &= \sum_{l=0}^{L-1} c_l \int_{-\infty}^{\infty} p(kT_{\mathrm{s}} - lT_{\mathrm{s}} - \tau)p^*(-\tau)e^{j(\theta(kT_{\mathrm{s}})-\tau)}\mathrm{d}\tau \\ &+ \int_{-\infty}^{\infty} \tilde{n}'(kT_{\mathrm{s}} - \tau)p^*(-\tau)\mathrm{d}\tau \\ &= c_k e^{j\theta(kT_{\mathrm{s}})} + \tilde{n}(kT_{\mathrm{s}}), \end{aligned} \tag{3.6}$$

where $k \in \mathbb{Z}^+$, $r(kT_{\mathrm{s}})$ is the received signal sample, $\tilde{n}(kT_{\mathrm{s}})$ is the complex Gaussian noise sample with $\mathbb{E}\{\tilde{n}(kT_{\mathrm{s}})\} = 0$, $\mathbb{E}\{\tilde{n}(kT_{\mathrm{s}})\tilde{n}^*(kT_{\mathrm{s}})\} = N_0$, and $\theta(kT_{\mathrm{s}})$ is the phase noise sample in the kth time instant. Here, $\mathbb{E}$ denotes the expectation operator. The simplification in (3.6) results because $p(t)$ is a square root Nyquist pulse, and it is assumed that *the phase noise variation is constant within* T_{s}[1]. The discrete phase noise process $\theta(kT_{\mathrm{s}})$ can be written using (2.2) and (2.4) as

$$\begin{aligned} \theta(kT_{\mathrm{s}}) &= \sum_{i=1}^{k} \int_{(i-1)T_{\mathrm{s}}}^{iT_s} \theta(t)\mathrm{d}t = \sum_{i=1}^{k} \Delta(iT_s) \\ &= \theta((k-1)T_{\mathrm{s}}) + \Delta(kT_{\mathrm{s}}). \end{aligned} \tag{3.7}$$

With a slight change in notation, we rewrite the discrete phase noise process in (3.7) as

$$\theta_k = \theta_{k-1} + \Delta_k, \tag{3.8}$$

where θ_0 is a uniform random variable (RV), and $\Delta_k \sim \mathcal{N}(0, \sigma_\Delta^2)$ is the innovation of the Wiener phase noise process. For the discrete Wiener process in (3.8), the innovation process is white and distributed as $\Delta_k \sim \mathcal{N}(0, \sigma_\Delta^2)$, where σ_Δ^2 is defined in (2.6) as $\sigma_\Delta^2 = 4\pi^2 K_{\mathrm{w}} T_{\mathrm{s}}$.

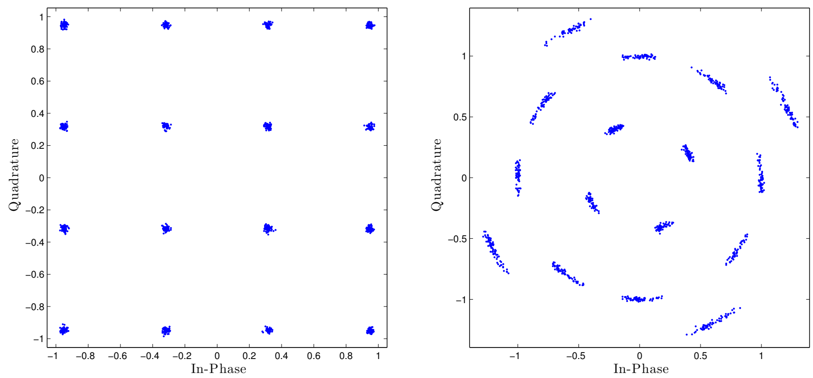
Finally, we rewrite the discrete system model in (3.6) as

$$r_k = c_k e^{j\theta_k} + \tilde{n}_k. \tag{3.9}$$

The discrete signal r_k in (3.9) forms a sufficient statistics for the continuous time model in (3.5) [22], under the assumption that the phase noise variation is constant within T_{s}. As we can see in (3.9), phase noise results in the random rotation of the transmitted information symbol, c_k. As an example, let us now visualize in Fig. 3.1 the effect of the Gaussian phase error on a 16-QAM constellation, where the signal-to-noise (SNR) per bit is 30 dB, and the innovation variance[2] is set to $\sigma_\Delta^2 = 10^{-4}$ rad^2.

[1] In [21], multi-sample receivers, where the received signal is sampled multiple times per symbol, have been considered upon relaxing this assumption.

[2] The innovation variance can be computed from the oscillator specifications [8].



(a) (b)

Figure 3.1: 16-QAM constellation at SNR per bit of 30 dB, when (a) no phase noise is present, and when (b) the signal is affected by the Wiener phase noise process with innovation variance $\sigma_{\Delta}^2 = 1 \times 10^{-4}$ rad^2.

For the remainder of the thesis, we will use the discrete model in (3.9) to represent an information signal that is affected by phase noise and AWGN in a SISO system. In order to recover c_k, receiver algorithms have to be designed to estimate θ_k, followed by the appropriate compensation of r_k, and the detection of c_k. But, before delving into the problem of receiver design for systems impaired by phase noise, it is important to understand the effect of channel fading on c_k, in addition to phase noise and AWGN.

3.2 SISO System Model with Phase Noise and Channel Fading

In (3.9), the effect of channel fading on the received signal has not been taken into account. In practice, the transmitted signal experiences channel fading, in addition to AWGN and phase noise. In particular, we are concerned about the effect of time-varying channel fading on the received signal, which depends on the relative velocity, v, between the transmitter and the receiver. In the presence of channel fading and phase noise [23, 24], the discrete-time system model in (3.9) can be rewritten as

$$r_k = h_k c_k e^{j\theta_k} + \tilde{n}_k. \tag{3.10}$$

Here, h_k denotes the channel gain between the transmitter and the receiver, and we assume that the channel fading process is based on the Clarke's model [24],

$$h_k \sim \mathcal{CN}(0, 1),\ \mathbb{E}\{h_k h_l^*\} = \mathrm{J}_0(2\pi f_{\mathrm{D}} T_{\mathrm{s}} |l - k|). \tag{3.11}$$

In (3.11), J_0 is the zero-order Bessel function of the first kind, and f_{D} is the maximum Doppler frequency given by

$$f_{\mathrm{D}} = \frac{v f_{\mathrm{osc}}}{c}, \tag{3.12}$$

where $c = 3 \times 10^8$ [m/s]. From (2.6), (3.10), and (3.12), it can be seen that the relative severity of phase noise and channel fading depends on f_{osc}, bandwidth, v, and the quality of the oscillators used in the

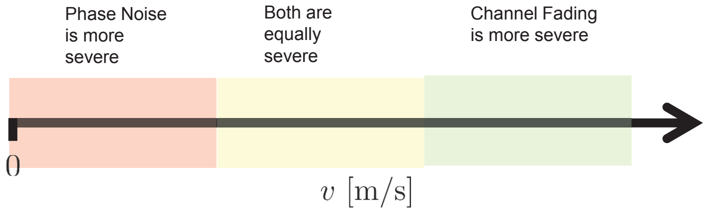
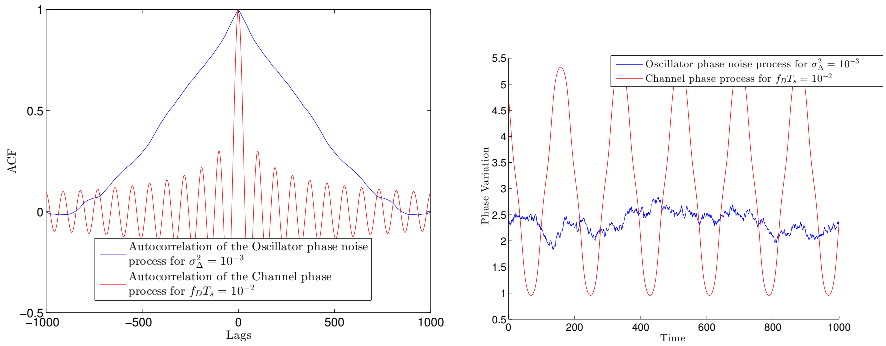


Figure 3.2: Regions where phase noise and channel fading are dominant in terms of their effects on the received signal, as a function of the relative velocity v. Note that the transition from one region to another depends on the quality of the oscillator.



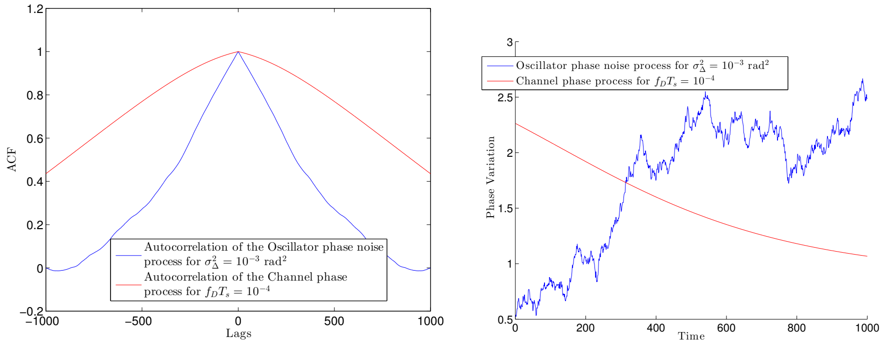
(a) (b)

Figure 3.3: Channel vs Phase Variations for $\sigma_\Delta^2 = 10^{-3}$ rad^2 or $\sigma_\Delta = 2^\circ$ and $f_\mathrm{D} T_\mathrm{s} = 10^{-2}$, (a) Correlation, (b) Phase variation.

transceivers. As shown in Fig. 3.2, when v is small, we can expect phase noise to be more dominant. However, as v increases, we expect the channel fading to be more dominant, and this scenario has been extensively studied in the literature [25]. Between these extreme scenarios, we anticipate a region where both phase noise and channel fading are equally dominant, which is discussed in Sec. 3.2.1.

In this thesis, we are particularly interested in the scenarios where the channel fading is less dominant than phase noise, i.e., the channel varies much slower than the phase noise, and can be considered to be quasi-static. These scenarios are elucidated in the following, based on which we motivate a system model where the channel is known or estimated accurately. First, we present the phase noise process, the channel phase process, and their respective autocorrelation functions for $\sigma_\Delta^2 \approx 2^\circ$ ($\sigma_\Delta^2 \approx 10^3$ rad^{-2}), $f_\mathrm{D} T_\mathrm{s} = 10^{-2}$ in Fig. 3.3, and $f_\mathrm{D} T_\mathrm{s} = 10^{-4}$ in Fig. 3.4. It can be seen in Fig. 3.3 that as the Doppler spread increases, the channel phase varies much faster than the phase noise. However, in low Doppler spread scenarios, the opposite becomes true. As can be seen in Fig. 3.4, phase noise tends to have large variations from one sample to the next, and varies much faster than the channel phase process. Furthermore, based on Fig. 3.4, it is possible to consider the channel process as a block fading process, given the slow-varying nature of the channel—i.e., it is possible to define transmission blocks that are small enough such that the channel is a blockwise constant.

Examples: Scenarios where the phase noise process changes much faster than the channel process



(a) (b)

Figure 3.4: Channel vs Phase Variations for $\sigma_\Delta^2 = 10^{-3}$ rad^2 or $\sigma_\Delta = 2°$ and $f_\mathrm{D}T_\mathrm{s} = 10^{-4}$, (a) Correlation, (b) Phase variation.

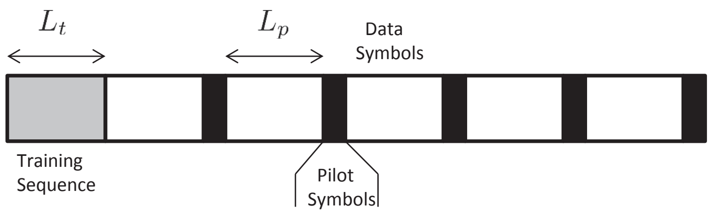


Figure 3.5: Frame Structure from [28]

typically occurs in microwave backhaul networks. In these networks, the channel essentially remains constant for a long period of time (quasi-static fading), and the phase noise is much more severe than the channel. This is also similar to the case of line-of-sight (LoS) MIMO systems, where a full-rank channel matrix is achieved by a careful placement of the antennas [26]. In this case, the channel is almost constant, and the phase noise is a major impairment. Lastly, as the carrier frequency increases, it is expected that the phase noise innovation variance will increase significantly, which will make the phase noise a limiting factor in millimeter-wave systems [27]. $\triangle$

Motivated by the aforementioned scenarios, in this thesis, we assume that the channel gains are estimated accurately at the receiver. In order to estimate the channel, a frame structure as in Fig. 3.5 is employed. Based on a training sequence of L_t symbols, joint channel and phase noise estimation is performed using a least-square (LS) estimator [28]. This is followed by the transmission of data symbols, embedded with a pilot symbol every L_p data symbols that are used for phase noise estimation [28, p. 4793]. The estimates obtained from the LS estimator are used as the true channel realizations in the phase noise estimation algorithm. Thus, we assume that h_k in (3.10) is known, and we design algorithms to estimate the phase noise, and further detect the transmitted data. This corresponds to the system model used in Papers A, B, and C.

3.2.1 Joint Models for Phase Noise and Channel Fading in SISO Systems

In this section, we take a detour by describing a system model where the transmitted signal is impaired by both phase noise and channel fading, and both the processes are equally severe (see Fig. 3.2). That is, here, both phase noise and the channel vary at comparable rates. In an attempt to simplify the problem description and the analysis, based on (3.10) we combine the phase noise and the channel variations as $g_k \triangleq h_k e^{j\theta_k}$. The joint channel-phase noise process g_k can then be approximated as a 1st order auto-regressive (AR) process,

$$g_k \approx \rho g_{k-1} + v_k, \tag{3.13}$$

where $v_k \sim \mathcal{CN}(0, \sigma_{\mathrm{v}}^2)$. The variance of v_k, which is denoted by σ_{v}^2, is calculated as

$$\begin{aligned}
\sigma_{\mathrm{v}}^2 &= \mathbb{E}(g_k - \rho g_{k-1})(g_k - \rho g_{k-1})^* & (3.14)\\
&= \mathbb{E}|g_k|^2 + \rho^2 \mathbb{E}|g_{k-1}|^2 - 2\rho \mathbb{E}\Re\{g_k g_{k-1}{}^*\} & (3.15)\\
&= 1 + \rho^2 - 2\rho \mathbb{E}\Re\{h_k h_{k-1}^* e^{j\theta_k - \theta_{k-1}}\} & (3.16)\\
&= 1 + \rho^2 - 2\rho \mathrm{J}_0(2\pi f_{\mathrm{D}} T_{\mathrm{s}}) e^{-\frac{\sigma_\Delta^2}{2}}. & (3.17)
\end{aligned}$$

We now find ρ that minimizes σ_{v}^2 by differentiating (3.17) with respect to ρ, and setting the result to zero. This yields

$$\begin{aligned}
\rho &= \mathrm{J}_0(2\pi f_{\mathrm{D}} T_{\mathrm{s}}) e^{-\frac{\sigma_\Delta^2}{2}} & (3.18)\\
&\Longrightarrow \sigma_{\mathrm{v}}^2 = 1 - \rho^2. & (3.19)
\end{aligned}$$

Based on (3.19), a joint channel-phase noise estimator can be designed using a Kalman Filter or the SPA. However, the problem of designing receivers based on the joint model described is not considered in this thesis, and is identified as future work in Chapter 6.

Till now, we focussed on a SISO system, which is considered in Papers A and B. Now, we shift our focus to multiple-antenna systems, and discuss about the impact of phase noise in these systems.

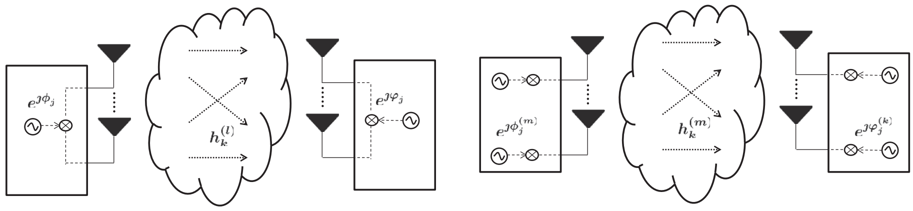
3.3 MIMO Systems

An avenue for increasing the data rate in a communication link is to use multiple antennas both at the transmitter and the receiver [29]. Such a system combined with transmission/reception techniques to exploit the degrees of freedom provided by the multi-path fading channel is referred to as a MIMO system. Beamforming is one technique that can be used along with multiple antennas in order to improve the receive SNR (array gain), and reliability (diversity gain). However, this technique requires the availability of reliable CSI. In the absence of reliable CSI, space-time coded transmission can be performed, which offers a diversity gain at the receiver, but does not improve the array gain. Compared to SISO systems, MIMO systems have higher spectral efficiency and diversity gain.

In the power-limited regime, the capacity of a MIMO system grows with the number of transmit antennas, N_{t}, and the receive antennas, N_{r}, as $C = \min(N_{\mathrm{t}}, N_{\mathrm{r}}) \log_2(\mathsf{SNR}) + \mathcal{O}(1)$ [18], where SNR denotes the SNR at the receiver. This is referred to as the spatial multiplexing gain. It is also possible to achieve spatial multiplexing gains in an LoS MIMO system. This is realized by a careful geometric placement of the antennas at the transmitter and the receiver.

3.3.1 MIMO System Model with Phase Noise

In general, the analysis and the design of a MIMO system is based on the assumption that the carrier phase is perfectly known at the receiver, and that there is no phase noise in the system. However, in the presence



(a) (b)

Figure 3.6: (a) The common oscillator setup, and (b) the distributed oscillator setup in point-to-point MIMO system.

of noisy local oscillators, phase noise results in a time-varying phase difference between the transmitter and the receiver. In MIMO systems, the antennas at the transmitter and the receiver can be connected to the oscillators in different ways. Two setups are particularly interesting, and they are shown in Fig. 3.6. In the first, a common oscillator is connected to all antennas at the transmitter/receiver (referred to as the common oscillator (CO) setup). In the second, each antennas has its own oscillator (referred to as the distributed oscillator (DO) setup). In the sequel, we will illustrate some aspects of the DO setup, which makes it inherently different from the CO setup.

The system model for the DO setup is as follows[3]. Data is transmitted as frames of symbols, and the channel between the transmit and receive antennas is assumed to be a constant over the length of a frame, as in the SISO case. Each antenna is equipped with an independent free-running oscillator that causes phase noise, which varies symbol by symbol, and much faster than the channel [30]. We assume that for a given frame, the channel and phase noise are first jointly estimated as discussed in Section 3.2.1. The training sequence is then followed by the transmission of data symbols, during which an autonomous phase noise estimator is used to track the phase noise, followed by a data detector. The joint channel and phase noise estimates obtained from the training sequence are used as the true channel values in the phase noise estimation and the data detection algorithm.

Assuming square-root Nyquist pulses for transmission, and matched filtering followed by sampling at symbol period T_{s}, the received signal in the kth time instant at the nth receive antenna is

$$\begin{aligned} r_k^{(n)} &= \sum_{m=1}^{N_{\mathrm{t}}} h_k^{(m,n)} c_k^{(m)} e^{\jmath(\phi_k^{(m)}+\varphi_k^{(n)})} + w_k^{(n)} \\ &\triangleq \sum_{m=1}^{N_{\mathrm{t}}} c_k^{(m,n)} e^{\jmath\theta_k^{(m,n)}} + w_k^{(n)}. \end{aligned} \tag{3.20}$$

In (3.20), $c_k^{(m)} \in \mathcal{M}$ is the symbol transmitted from the mth transmit antenna at the kth time instant, and drawn equiprobably from a C-ary signal constellation set $\mathcal{M}$, $h_k^{(m,n)}$ represents the known (or estimated) channel realization between the mth transmit and nth receive antenna, $c_k^{(m,n)} \triangleq c_k^{(m)} h_k^{(m,n)}$ and $w_k^{(n)} \sim \mathcal{CN}(0, N_0)$ denotes the zero-mean AWGN at the nth receive antenna. The phase noise at time instant k in the (m,n)th link is $\theta_k^{(m,n)} \triangleq \phi_k^{(m)} + \varphi_k^{(n)}$, where $\phi_k^{(m)}$ and $\varphi_k^{(n)}$ denote the oscillator phase noise sample at the mth transmit and the nth receive antennas, respectively.

Traditionally, for the system model in (3.20), MIMO receiver design has focused on developing algorithms for joint channel estimation and data detection (we refer the readers to [25, 31] and the references

[3]The received signal model for the DO setup can be trivially reduced to that of the CO setup.

therein). It is perceived that the phase noise can be handled by existing channel estimation-data detection algorithms, since the phase noise can be treated to be part of the channel [23]. However, for the scenarios discussed in 3.2.1, phase noise cannot be treated to be a part of the channel, and has to be compensated separately. Thus, joint phase noise estimation and data detection algorithms have to be designed by assuming that the channel is known (estimated), and that the channel varies much slower than the phase noise process.

For the DO setup, it is worth noting that the actual number of phase variables to be estimated can be reduced to $N_{\mathrm{t}} + N_{\mathrm{r}} - 1$, as opposed to estimating $N_{\mathrm{t}} N_{\mathrm{r}}$ variables (in $N_{\mathrm{t}} N_{\mathrm{r}}$ links). This is made possible by subtracting all the transmit phases by any one of the transmit phases and adding the same amount to all the receive antenna phases. For instance, in (3.20) the phase states to be estimated are $\{\theta_k^{(m,n)}\}$, $m = 1, \ldots, N_{\mathrm{t}}$, $n = 1, \ldots, N_{\mathrm{r}}$, implying that there are $N_{\mathrm{t}} N_{\mathrm{r}}$ phase noise variables to be estimated. However, the transmit phase noise variables can be transformed to $\{0, \phi_k^{(m)} - \phi_k^{(1)}\}$, $m = 2, \ldots, N_{\mathrm{t}}$, and the receive phase noise variables can be changed to $\{\varphi_k^{(n)} + \phi_k^{(1)}\}$, $n = 1, \ldots, N_{\mathrm{r}}$. This transformation, in effect, produces the same received signal model in (3.20), even though the transmit and receive phase states to be estimated have been altered, and reduced to $N_{\mathrm{t}} + N_{\mathrm{r}} - 1$ states.

Furthermore, in contrast to a SISO system or a MIMO system with the CO setup, it is interesting to see that connecting each of the antennas to a different oscillator gives rise to *both phase distortions and amplitude distortions* in the received signal. This can be illustrated using a simple example:- for a 2×2 MIMO system, the received signal at receive antenna $n = 1$ at high SNR can be written as

$$\begin{aligned} r_k^{(1)} &\approx e^{\jmath\theta_k^{1,1}} h_k^{(1,1)} c_k^{(1)} + e^{\jmath\theta_k^{(2,1)}} h_k^{(2,1)} c_k^{(2)}, \\ r_k^{(1)} r_k^{(1)^*} = |r_k^{(1)}|^2 &= c_k^{(1)} c_k^{(1)^*} h_k^{(1,1)} h_k^{(1,1)^*} + c_k^{(2)} c_k^{(2)} h_k^{(2,1)} h_k^{(2,1)^*} \\ &\quad + 2\Re\{c_k^{(1)} c_k^{(2)^*} h_k^{(1,1)} h_k^{(2,1)^*} e^{\jmath(\theta_k^{(1,1)} - \theta_k^{(2,1)})}\}. \end{aligned} \tag{3.21}$$

As evident from (3.21), the amplitude of the received signal depends on the phase difference between the signals arriving at the receive antenna. In addition, it is shown that phase noise in the DO setup can cause relatively more severe estimation errors in the capacity for a MIMO link. This results from the incorrect calculation of the channel rank from channel sounding experiments [23]. Receiver algorithm design for the CO setup is similar to that for a SISO system. However, in the DO setup, new algorithms based on the maximum a posteriori (MAP) detection theory have to be derived. This forms the crux of our contribution in Paper C [6].

The performance of point-to-point MIMO systems depends on the availability of a rich scattering environment. In the absence of such an environment, the system performance experiences severe degradation. This shortcoming is overcome by considering a multiuser (MU) MIMO system, which is explained in the following section.

3.4 Multiuser MIMO and Massive MIMO Systems

The linear growth of the capacity with $\min(N_{\mathrm{t}}, N_{\mathrm{r}})$ for a point-to-point MIMO link is achievable only in the presence of sufficient scatterers—in the absence of scatterers, the channel matrix becomes rank-deficient causing the spatial multiplexing gains to vanish. In a cellular system, it is possible to have multiple antennas at the base station (BS), while the users may have a relatively lower number of antennas, due to size and cost constraints. Therefore, the capacity of the point-to-point link is limited by the number of antennas at the users. An alternative is to consider a multi-user MIMO (MU-MIMO) system [32, 33], which uses MU transmission and reception methods. Here, spatial diversity that arises from the geographical separation between the users, irrespective of the number of antennas at the user, helps to realize the array and diversity gain achievable in a point-to-point MIMO system. Some key advantages of a MU-MIMO system are as follows.

- The MU multiplexing gain can be realized without multiple antennas at the users, thereby allowing the development of small and cheap terminals.

- The system performance is more resilient to the limitations imposed by the propagation environment and channel correlations.

The benefits of a MU-MIMO system have triggered tremendous interest in the area of massive MIMO, which is a MU-MIMO system, where the BS has a large number of antennas. This is discussed in the sequel.

3.4.1 Potential and Challenges of Massive MIMO

Massive MIMO is considered to be a key enabler for the development of future broadband wireless networks [18, 19]. It is envisioned that a massive MIMO system will comprise of BSs that use antenna arrays consisting of several hundreds or thousands of antennas. These BSs are expected to serve multiple user equipments (UEs) (with single antenna) in the same time-frequency resource blocks. That is, a massive MIMO system is a MU-MIMO system with a large number of BS antennas. The cardinal aspect of a massive MIMO system is that the channel vectors between the UEs and the BS antennas are pairwise orthogonal, which is referred to as the favorable propagation (f.p.) condition. The f.p. condition arises due to the spatial separation of the multiple UEs, and the fact that the channel vectors are asymptotically long due to a large number of BS antennas. The main benefits of massive MIMO are as follows.

- Massive MIMO can significantly increase capacity. This arises from the MU multiplexing gain due to MU transmission and reception methods.

- The radiated energy efficiency increases dramatically. This stems from the fact that as the antenna aperture increases, energy can be focussed into small spatial regions.

- Massive MIMO can be built with inexpensive, low-power components. In effect, massive MIMO is expected to reduce the constraints on the quality of the individual components, since the massiveness will average out some of the impairments, which results from using nonideal hardware components in the transceiver.

- Massive MIMO facilitates reduction in latency, and also simplifies the multiple access layer due to the averaging of the small-scale fading channels (channel hardening).

- The f.p. condition facilitates the use of simple linear transmission and reception methods, which have close-to-optimum performance.

Benefits of a massive MIMO system can be reaped only in the presence of reliable CSI at the BS for both uplink and downlink transmissions. In conventional MIMO systems, the BS transmits pilot signals, which are used by the UEs to estimate the CSI, and this information is fed back to the BS. Such a channel training scheme is not feasible in a massive MIMO system given the large number of BS antennas and the finite channel coherence time. Hence, it is expected that a massive MIMO system will benefit from a time division duplex (TDD) mode by exploiting the reciprocity between the uplink and the downlink channels [19]. In addition to acquiring reliable CSI, the other challenges in a massive MIMO are the following [18, 19]:

- TDD operation relies on channel reciprocity. However, the difference in the transceiver chain at the UE and the BS results in non-reciprocity between the uplink and downlink channels.

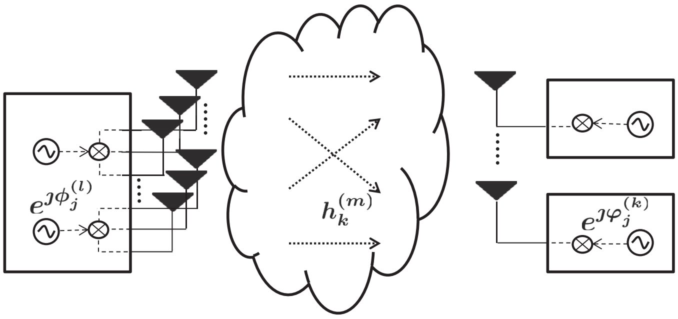


Figure 3.7: The general oscillator (GO) setup, where the BS has M_{osc} free-running oscillators, and $M/M_{\text{osc}} \in \mathbb{Z}^+$ BS antennas are connected to each oscillator

- Ideally, the uplink pilot signals assigned to all the UEs in the network should be orthogonal. This cannot be realized in practice, and thus non-orthogonal pilots are used across the cells in the network. This results in pilot contamination, where the estimate of the channel between a UE and the BS is contaminated by the non-orthogonal pilots transmitted concurrently by the other UEs.

- Massive MIMO is expected to be constructed using low-cost, and low-energy hardware [34]. This is on the pretext of the law of large numbers, which is expected to average out noise due to imperfect hardware. Some of the impairments include high levels of quantization noise that can get injected into the transceiver due to (low power) low resolution A/D converters. Low-cost phase locked-loops or free-running oscillators on each RF chain feeding an antenna at the BS can result in severe phase noise impairment, which can significantly affect system performance.

3.4.2 Massive MIMO and Phase Noise

As discussed before, it is well understood that the performance of massive MIMO systems can be severely limited by impairments arising from the non-ideal transceiver hardware components [34]. Implementing transceiver algorithms mandates the availability of reliable CSI at the BS [18]. This is challenging as the coherence time of the channels between the BS and its associated UEs is finite, and thus the BS is required to update its CSI regularly. Furthermore, hardware impairments affect the CSI quality drastically, and the problem exacerbates in the presence of phase noise due to noisy local oscillators [9, 34, 35].

In massive MIMO systems, phase noise causes random rotation of the transmitted data symbols. Furthermore, as seen in [9, 35], phase noise causes the actual effective channel phase during the data transmission period to become significantly different from that during the training period. This is because the effective channel phase drifts randomly between the time instant that a pilot symbol is received and the instant that a data symbol is transmitted/received. This is referred to as the *channel-aging phenomenon* [36]. In this light, it becomes important to analyze the impact of phase noise on the performance of massive MIMO systems.

In this thesis, we investigate the uplink and downlink performance of a massive MIMO system in the presence of phase noise. In Paper D [8], we consider the downlink channel where we analyze linear precoding schemes. We consider a single-cell massive MIMO system comprising of one BS with M antennas serving K single-antenna UEs. We analyze a general oscillator (GO) setup, shown in Fig. 3.7, where the BS has M_{osc} free-running oscillators, and $M/M_{\text{osc}} \in \mathbb{Z}^+$ BS antennas are connected to each oscillator. The CO and DO setups, which are discussed earlier, are special cases of this general setup. In Paper E [9], we analyze the impact of phase noise due to noisy local oscillators for a massive MIMO system that consists of a BS and a single-antenna UE, and orthogonal-time division multiplexing transmission (OFDM) is considered [37]. We study the effect of channel-aging in the CO and the DO setups, and we analyze the SNR on each subcarrier for $M \to \infty$, when a maximum-ratio combining (MRC) receiver is used. In order to conduct our analysis for both uplink and downlink transmissions, we use tools from random matrix theory (RMT) and free probability [38]. Relevant mathematical preliminaries are provided in Chapter 4.

Chapter 4

Tools for Analysis and Design of Communication Systems with Phase Noise

This chapter is an assortment of the various tools and results used in this thesis for analyzing and designing communication systems in the presence of phase noise. First, we discuss about Bayesian inference methods, and their application to derive the MAP symbol detector for the SISO phase noise channel. This is followed by a brief description of the sum-product algorithn (SPA), and the variational Bayesian technique. These tools are employed in Papers A, B, and C. Then, we discuss about results related to the Shannon capacity of the SISO phase noise channel. Finally, we present a short tutorial on random matrix theory (RMT). RMT is used to analyze the performance of massive MIMO systems in the presence of phase noise in Papers D and E.

4.1 Bayesian Inference Methods and Their Applications to SISO Systems with Phase Noise

Bayesian inference methods use Bayes' rule to update the probability estimates of different hypotheses in an experiment [39, 40]. Updating the probability estimate of a hypothesis is particularly important when analyzing data. In communication systems, Bayesian inference methods are useful for analyzing the noisy received signals. In particular, they can be used for estimating nuisance parameters, such as phase noise or channel coefficients, and detecting the transmitted information [41, 42]. One of the early works that studied the problem of optimal signal detector in a Bayesian setting is [43]. In this work, a SISO channel with AWGN is considered, and the optimal signal detector is determined such that the detection rate is maximized for a given false alarm rate (i.e., the chance for pure noise to be detected as information). This approach has been extended to more realistic communication systems in [44]. In recent times, the popularity of designing receivers based on the MAP theory has increased with the development of powerful, low-complexity algorithm frameworks like the SPA, variational Bayesian method, graphical method etc. [45].

In the following, we study the optimal receiver, which minimizes symbol error probability (SEP) for uncoded data transmission in the presence of oscillator phase noise [46]. The optimal receiver is realized by the MAP symbol detector. Furthermore, we study the basics of factor graphs (FGs), and the variational Bayesian frameworks, which helps to realize the optimal receiver algorithm.

4.1.1 MAP Symbol Detector

In this section, we discuss the MAP symbol detector derived in [46] for the SISO phase noise channel. For the system model in (3.6), the MAP symbol detector is written as

$$\begin{aligned}\hat{c}_k &= \arg\max_{c_k} \sum_{\boldsymbol{c}\setminus\{c_k\}} P(\boldsymbol{c}|\boldsymbol{r}) &(4.1)\\ &\propto \arg\max_{c_k} \int_{\theta_k} P(c_k)p(r_k|c_k,\theta_k)p(\theta_k|\bar{\boldsymbol{r}}_k)\mathrm{d}\theta_k, &(4.2)\end{aligned}$$

where $\boldsymbol{c} = [c_1, \ldots, c_L], \bar{\boldsymbol{r}}_k = [r_1, \ldots, r_{k-1}, r_{k+1}, \ldots, r_L]$, L is the number of symbols transmitted, and $\setminus$ denotes a set-subtract.

In uncoded systems, the MAP detector that minimizes the SEP for uncoded data transmission determines the transmitted symbols based on (4.2). In this receiver, the symbols are detected based on the a posteriori pdf of the phase noise process conditioned on $\bar{\boldsymbol{r}}_k$. This is in contrast to many estimator-detector structures studied in the literature [46], where the symbol detection is performed based on the MAP estimate of the phase noise process. For the Wiener phase noise process, determining $p(\theta_k|\bar{\boldsymbol{r}}_k)$ is analytically intractable, which also makes the MAP detector intractable and unimplementable in practice [46]. Thus, it is imperative to make approximations of the MAP symbol detector. We consider two Bayesian approaches for performing approximations, namely, the SPA based on factor graphs [45], and the variational Bayesian algorithm [47].

4.1.2 Factor Graphs and the Sum Product Algorithm

In this section, we introduce FGs and the SPA by deriving the MAP symbol detector in (4.2) for the system model in (3.6)[1]. Broadly, FGs belong to the family of graphical models, which are used to represent factorization of multivariate functions—they can be used to visualize interaction between the variables of a function. In order to represent a multivariate function in the form of an FG, we express the actual (global) function as a product of simpler (local) functions, each of which depends only on a subset of variables. The graphical representation of the factorized function yields a bipartite graph, which expresses the dependencies between the variables, and the local functions. FGs combined with SPA can be used for computing marginals of probability functions. In our problem setting, marginalization of the phase noise process is an important step towards realizing the MAP symbol detector.

In order to derive the MAP detector using the SPA, we rewrite (4.1) as

$$\begin{aligned}\hat{c}_k &= \arg\max_{c_k} \sum_{\boldsymbol{c}\setminus\{c_k\}} P(\boldsymbol{c}|\boldsymbol{r})\\ &= \arg\max_{c_k} \sum_{\boldsymbol{c}\setminus\{c_k\}} \int_{\boldsymbol{\theta}} P(\boldsymbol{c},\boldsymbol{\theta}|\boldsymbol{r})\mathrm{d}\boldsymbol{\theta}. &(4.3)\end{aligned}$$

Factorizing the integrand, we obtain

$$\begin{aligned}P(\boldsymbol{c},\boldsymbol{\theta}|\boldsymbol{r}) &\propto P(\boldsymbol{c})p(\boldsymbol{\theta}|\boldsymbol{c})p(\boldsymbol{r}|\boldsymbol{c},\boldsymbol{\theta}),\\ &= P(\theta_0)\prod_{k=1}^{L} P(c_k)\underbrace{p(\theta_k|\theta_{k-1})}_{p_\Delta(\theta_k-\theta_{k-1})} p(r_k|\theta_k,c_k). &(4.4)\end{aligned}$$

[1] For a basic tutorial on FGs and the sum product algorithm, we refer the readers to [45].

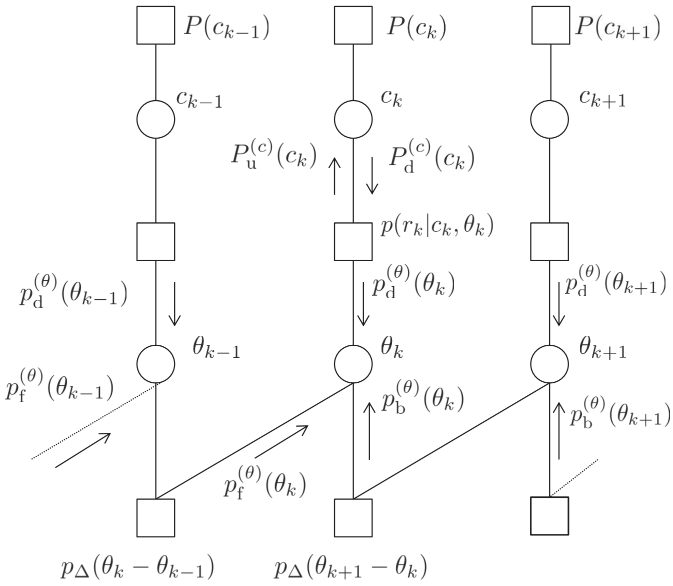


Figure 4.1: Factor Graph and the SPA messages based on (4.4)

To factorize the function in (4.4) we exploit the fact that θ_k is a discrete Wiener process as in (3.8). The FG associated with the global function in (4.4) is shown in Fig. 4.1. The messages on the graph are

$$
\begin{aligned}
P_{\mathrm{d}}^{(c)}(c_k) &= P(c_k) && (4.5)\\
p_{\mathrm{d}}^{(\theta)}(\theta_k) &= \sum_{c_k} P_{\mathrm{d}}^{(c)}(c_k) p(r_k|c_k, \theta_k) && (4.6)\\
p_{\mathrm{f}}^{(\theta)}(\theta_k) &= \int_{\theta_{k-1}} p_{\mathrm{f}}^{(\theta)}(\theta_{k-1}) p_{\mathrm{d}}^{(\theta)}(\theta_{k-1}) \cdot p_\Delta(\theta_k - \theta_{k-1}) \mathrm{d}\theta_{k-1} && (4.7)\\
p_{\mathrm{b}}^{(\theta)}(\theta_k) &= \int_{\theta_{k+1}} p_{\mathrm{b}}^{(\theta)}(\theta_{k+1}) p_{\mathrm{d}}^{(\theta)}(\theta_{k+1}) p_\Delta(\theta_{k+1} - \theta_k) \mathrm{d}\theta_{k+1} && (4.8)\\
P_{\mathrm{u}}^{(c)}(c_k) &= \int_{\theta_k} p_{\mathrm{f}}^{(\theta)}(\theta_k) p_{\mathrm{b}}^{(\theta)}(\theta_k) p(r_k|c_k, \theta_k) \mathrm{d}\theta_k. && (4.9)
\end{aligned}
$$

As can be seen in Fig. 4.1, $P_{\mathrm{d}}^{(c)}(c_k)$ in (4.5) is the message from variable node c_k towards the factor node $p(r_k|\theta_k, c_k)$. In (4.6), $p_{\mathrm{d}}^{(\theta)}(\theta_k)$ is the message from the factor node $p(r_k|\theta_k, c_k)$ towards the variable node θ_k. $p_{\mathrm{f}}^{(\theta)}(\theta_k)$ in (4.7) and $p_{\mathrm{b}}^{(\theta)}(\theta_k)$ in (4.8) are the messages from the factor nodes $p_\Delta(\theta_k - \theta_{k-1})$ and $p_\Delta(\theta_{k+1} - \theta_k)$, respectively, towards θ_k. Finally, $P_{\mathrm{u}}^{(c)}(c_k)$ in (4.9) is the message from the factor node $p(r_k|\theta_k, c_k)$ towards the variable node c_k.

Note that the FG in Fig. 4.1 is a tree, and thus applying the SPA on this graph gives the exact MAP symbol detector (4.2). Hence, $p_{\mathrm{b}}^{(\theta)}(\theta_k) p_{\mathrm{f}}^{(\theta)}(\theta_k)$ is the a posteriori pdf $p(\theta_k|\bar{\boldsymbol{r}}_k)$ given in (4.2). Thus, the detector in (4.2) can also be expressed as

$$\begin{aligned}\hat{c}_k &= \arg\max_{c_k} P_{\mathrm{u}}^{(c)}(c_k) P_{\mathrm{d}}^{(c)}(c_k) \\ &\propto \arg\max_{c_k} P_{\mathrm{u}}^{(c)}(c_k).\end{aligned} \tag{4.10}$$

Here, $P_{\mathrm{d}}^{(c)}(c_k)$ (which is also equal to $P(c_k)$) is uniform for uncoded transmission[2]. The messages in (4.5)-(4.9) form the core for the SPA-based implementation of the MAP detector. However, the implementation of the exact SPA is impractical because it involves the computation of the continuous pdfs of θ_k in (4.5)-(4.8), which are analytically intractable. The intractability of the exact MAP symbol detector in (4.2) and (4.10) motivates the need to explore practical, low complexity receiver algorithms based on the cannonical distribution approach in the SPA [49]. This approach involves constraining the messages on the FG to a specific family of pdfs, which can compactly and completely be described by a finite number of parameters. Thus, the task of computing the exact pdf is reduced to computing the parameters of the pdf. For instance, when the messages on the FG are constrained to belong to the exponential family of pdfs, then it suffices to determine the mean and variance to completely describe the pdf.

4.1.3 Variational Bayesian Framework

An alternative for realizing the optimal receiver in (4.2) is to use an algorithm that iteratively performs joint Bayesian estimation and symbol a posteriori probability computation. This can be realized by applying the variational Bayesian (VB) technique, which has been widely used by the communication engineering community for deriving efficient receiver algorithms, when the received signal is corrupted by random nuisance parameters [50]. A tutorial on the basics of this framework is available in [47], and in the sequel, we demonstrate the application of the VB framework for approximating the receiver in (4.2).

We first compute the log of the a prior probability of $\boldsymbol{r}$ as

$$\begin{aligned}\log p(\boldsymbol{r}) &= \log \sum_{\boldsymbol{c}} \int_{\boldsymbol{\theta}} p(\boldsymbol{c}, \boldsymbol{\theta}, \boldsymbol{r}) \mathrm{d}\boldsymbol{\theta} \\ &= \log \sum_{\boldsymbol{c}} \int_{\boldsymbol{\theta}} Q(\boldsymbol{c}, \boldsymbol{\theta}) \frac{p(\boldsymbol{c}, \boldsymbol{\theta}, \boldsymbol{r})}{Q(\boldsymbol{c}, \boldsymbol{\theta})} \mathrm{d}\boldsymbol{\theta} \\ &\geq \sum_{\boldsymbol{c}} \int_{\boldsymbol{\theta}} Q(\boldsymbol{c}, \boldsymbol{\theta}) \log \frac{p(\boldsymbol{c}, \boldsymbol{\theta}, \boldsymbol{r})}{Q(\boldsymbol{c}, \boldsymbol{\theta})} \mathrm{d}\boldsymbol{\theta}.\end{aligned} \tag{4.11}$$

When the variational distribution $Q(\boldsymbol{c}, \boldsymbol{\theta})$ is set to $P(\boldsymbol{c}, \boldsymbol{\theta}|\mathbf{r})$, the lower bound in (4.11) is achieved. However, the algorithm is restricted to search over a family of factorized distributions of the form: $Q(\boldsymbol{c}, \boldsymbol{\theta}) = Q(\boldsymbol{c}, \boldsymbol{\theta}) = q_{\boldsymbol{c}}(\boldsymbol{c}) q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. This corresponds to assuming that $\boldsymbol{c}$ and $\boldsymbol{\theta}$ are independent of each other given $\mathbf{r}$. Hence the lower bound is given by

$$\begin{aligned}\log P(\mathbf{r}) &\geq \sum_{\boldsymbol{c}} \int_{\theta} q_{\boldsymbol{c}}(\boldsymbol{c}) q_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \log \frac{P(\boldsymbol{c}, \boldsymbol{\theta}, \mathbf{r})}{q_{\boldsymbol{c}}(\boldsymbol{c}) q_{\boldsymbol{\theta}}(\boldsymbol{\theta})} d\boldsymbol{\theta}, \\ &\triangleq \mathcal{H}(q_{\boldsymbol{c}}(\boldsymbol{c}), q_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \mathbf{r}).\end{aligned} \tag{4.12}$$

Here, $\mathcal{H}(q_{\boldsymbol{c}}(\boldsymbol{c}), q_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \mathbf{r})$ is referred to as the inverse Gibbs or variational free energy, whose maximization results in the minimization of the *Kullback-Leibler* (KL) divergence measure between $q_{\boldsymbol{c}}(\boldsymbol{c}) q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$

[2] In coded systems, the same messages in (4.5)-(4.9) will be used, but $P_{\mathrm{d}}^{(c)}(c_k)$ is not the a priori pmf $P(c_k)$, but rather the extrinsic symbol pmf provided by the decoder. Also, the message $P_{\mathrm{u}}^{(c)}(c_k)$ in (4.9) is used for computing the bit log-likelihood ratios (LLRs) for soft decoding [5, 48].

and $P(\boldsymbol{c}, \boldsymbol{\theta}|\mathbf{r})$. In order to determine the factorized free distributions $q_{\boldsymbol{c}}(\boldsymbol{c})$ and $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ that maximize $\mathcal{H}$, a coordinate ascent algorithm is used that alternatingly maximizes over one free distribution while keeping the other fixed. Based on the functional derivatives of $\mathcal{H}$ with respect to the free distributions, the update equations are given as

$$\begin{aligned} q_{\boldsymbol{\theta}}(\boldsymbol{\theta}) &\propto P(\boldsymbol{\theta}) e^{\sum_{\boldsymbol{c}} q_{\boldsymbol{c}}(\boldsymbol{c}) \ln P(\mathbf{r}|\boldsymbol{c},\boldsymbol{\theta})}, \\ q_{\boldsymbol{c}}(\boldsymbol{c}) &\propto P(\boldsymbol{c}) e^{\int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \ln P(\mathbf{r}|\boldsymbol{c},\boldsymbol{\theta}) d\boldsymbol{\theta}}. \end{aligned} \tag{4.13}$$

The coordinate ascent algorithm converges to a fixed point [47], and in general global optimality is not guaranteed.

4.2 Capacity of Phase Noise Channels

A fundamental way to analyze the impact of random phase noise on the performance of a communication system is to determine the Shannon capacity. In [51], bounds for the capacity for a SISO system with uniform phase noise are derived. It is also shown that the capacity achieving pdf is discrete with infinite mass points. A similar conjecture is presented for partially coherent channels. In [52], the capacity achieving input pdf for partially coherent channels is found to be circularly symmetric, but not necessarily Gaussian distributed. In [53], upper bounds on the capacity for phase noise channels with and without memory are derived. Specifically, for the SISO channel given in (3.10), an upper bound for the achievable rate is given as

$$C_{\mathrm{PN}} = \min\{C_{1,\mathrm{PN}}, C_{2,\mathrm{PN}}\}, \tag{4.14}$$

where

$$C_{1,\mathrm{PN}} \leq \log_2\left(1 + \frac{|h|^2}{\sigma_{\mathrm{w}}^2}\right) \tag{4.15}$$

$$C_{2,\mathrm{PN}} \leq \frac{1}{2}\log_2\left(\frac{2\pi|h|^2}{\sigma_{\mathrm{w}}^2}\right) - \frac{1}{2}\log_2\left(2\pi e \tau(\sigma_{\varphi}^2 + \delta_{\mathrm{pn}}\sigma_{\phi}^2)\right) \tag{4.16}$$

In (4.16), the second term represents the differential entropy of the phase noise process, $\varphi_k + \phi_k$. The result in (4.16) holds under the assumption that the phase-noise process is stationary, and has a finite differential-entropy rate. The results in (4.14), (4.15) and (4.16) are used in Papers D and E in order to analyze the performance of the massive MIMO systems. For a survey of capacity results relevant to MIMO systems, refer to [54].

4.3 Random Matrix Theory and Asymptotic Results

RMT is widely applied to problems in physics, statistics, data analysis and engineering [38, 55]. In the last few years, a large body of work has emerged in the field of communications and information theory that have not only employed RMT results, but also have made fundamental contributions to RMT [56]. Tools from RMT have been particularly attractive to researchers for analyzing the performance of massive MIMO systems, where typically the analysis involves random matrices of large dimensions.

Consider a random matrix denoted by $\boldsymbol{H}$ of size $M \times K$, whose entries are Gaussian i.i.d. RVs. Specifically, the element in the ith row and the jth column in $\boldsymbol{H}$ is denoted by $\boldsymbol{H}|_{i,j} \sim \mathcal{CN}(0, \frac{1}{M})$. In a massive MIMO system, $\boldsymbol{H}$ can represent the small-scale Rayleigh fading channel matrix between K users and M BS antennas. As the number of rows and columns in $\boldsymbol{H}$ grows, i.e., $M, K \to \infty$, while $M/K = \beta$, the empirical cumulative distribution function of the eigenvalues (also called the spectrum) of

$\boldsymbol{H}$ shows interesting convergence properties. Specifically, the spectrum of $\boldsymbol{H}$ and its functionals become deterministic in the asymptotic limit. This observation leads to the central notion in asymptotic random matrix theory that the empirical distribution of the moments of the eigenvalues of $\boldsymbol{H}$ and its functionals become deterministic, and this is independent of the distribution of the matrix entries. Specifically, the spectrum of $\boldsymbol{H}\boldsymbol{H}^{\mathrm{H}}$ converges almost surely to a non-random distribution function called the Marchenko-Pastur law. These results are particularly useful given that eigenvalues of random matrices are used to characterize the performance of communication links (for e.g., MIMO links). Note that the nth moment of the eigenvalues of $\boldsymbol{H}$ is calculated as

$$\frac{1}{M}\sum_{m=1}^{M}\lambda_m^n = \frac{1}{M}\mathrm{tr}\{\boldsymbol{H}^n\}, \tag{4.17}$$

where λ_m denotes an eigenvalue of $\boldsymbol{H}$. This implies that the normalized trace of the functionals of $\boldsymbol{H}$, $\frac{1}{M}\mathrm{tr}\{\boldsymbol{H}^n\}$, becomes deterministic in the asymptotic limit. Even though the convergence of the spectrum is based on the assumption that both M and K become asymptotically large, this result is a good approximation even for small dimensions of $\boldsymbol{H}$ [57].

4.3.1 Stieltjes Transform

For a wide class of random matrices, the asymptotic eigenvalue distributions are either explicitly known or can be calculated numerically. However, the problem of determining an unknown probability distribution given its moments is addressed using the Stieltjes Transform.

DEFINITION 1 Stieltjes Transform [56, Section 2.2]: *Let X be a real-valued RV with distribution F. Then the Stieltjes transform $m(z)$ of F, for $z \in \mathbb{C}$ such that $\Im\{z\} > 0$, is defined as*

$$\begin{aligned} m(z) &= \mathbb{E}\left[\frac{1}{X-z}\right] = \int_{-\infty}^{\infty}\frac{1}{x-z}dF(x) && (4.18)\\ &= -\frac{1}{z}\sum_{n=1}^{\infty}\frac{\mathbb{E}[X^n]}{z^n}. && (4.19)\end{aligned}$$

The pdf of X, $p(x)$, can be obtained by invoking the Stieltjes inversion formula, which is given as

$$p(x) = \lim_{\omega\to 0^+}\frac{1}{\pi}m(x+\jmath\omega). \tag{4.20}$$

Based on (4.17), the Stieltjes transform can be viewed as the moment generating function of a random Hermitian matrix whose empirical eigenvalue distribution is $p(X)$.

4.3.2 Free Probability and Asymptotic Freeness

An important concept in asymptotic RMT analysis is that of non-commutative free probability theory [55, Chap. 22]. In free probability theory, a random matrix is viewed as a linear random operator, which is non-commutative, and the notion of statistical independence of random variables is overidden by that of "free independence" of random matrices. Consider the RVs, X and Y, and the expectation operator $\mathbb{E}$. Then,

$$\mathbb{E}(XY)^m = \mathbb{E}X^mY^m = \mathbb{E}X^m\mathbb{E}Y^m, \tag{4.21}$$

if x and y are statistically independent of each other. Now consider the random matrices (operators) $\boldsymbol{X}$ and $\boldsymbol{Y}$ of size $M \times M$, and the expectation operator for random matrices given as

$$\frac{1}{M}\mathrm{tr}(\boldsymbol{X}\boldsymbol{Y})^m \neq \frac{1}{M}\mathrm{tr}(\boldsymbol{X})^m(\boldsymbol{Y})^m \neq \frac{1}{M}\mathrm{tr}(\boldsymbol{X})^m \frac{1}{M}\mathrm{tr}(\boldsymbol{Y})^m, \tag{4.22}$$

even if the entries of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are statistically independent of each other, for $M \to \infty$. This is due to the non-commutative nature of matrix multiplication. In order to analyze the expectation operations in (4.22), free probability theory and the concept of free asymptotic independence have to be employed. Denote the expectation operator as $\mathrm{Tr}(\cdot) = \frac{1}{M}\mathrm{tr}(\cdot)$. Then, the matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are asymptotically freely independent from each other if

$$\mathrm{Tr}\left((P_X(\boldsymbol{X}) - \mathrm{Tr}P_X(\boldsymbol{X}))(P_Y(\boldsymbol{Y}) - \mathrm{Tr}P_Y(\boldsymbol{Y}))\right) \overset{M\to\infty}{\longrightarrow} 0 \tag{4.23}$$

where $P_X(\boldsymbol{X})$ and $P_Y(\boldsymbol{Y})$ are polynomials in $\boldsymbol{X}$ and $\boldsymbol{Y}$. In this thesis, we use free probability in order to simplify the algebra involving random matrices. Of particular interest is Lemma 1, which is given as follows.

LEMMA 1 *Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{C}^{M\times M}$ be freely independent random matrices with uniformly bounded spectral norm for all M [38, Page 207]. Further, let all the moments of the entries of $\boldsymbol{X}, \boldsymbol{Y}$ be finite, then,*

$$\mathrm{Tr}\boldsymbol{X}\boldsymbol{Y} - \mathrm{Tr}\boldsymbol{X}\mathrm{Tr}\boldsymbol{Y} \overset{M\to\infty}{\longrightarrow} 0. \tag{4.24}$$

In general, establishing free independence between random matrices is a non trivial problem. However, several interesting random matrices have been shown to be asymptotically free under certain conditions [56]. In Paper E, we use the following lemma to prove prove independence.

LEMMA 2 *Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{C}^{M\times M}$ be random matrices such that their asymptotic spectrum exists for $M \to \infty$ [38, Page 207]; the entries of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are statistically independent, and either $\boldsymbol{X}$ or $\boldsymbol{Y}$ are unitarily invariant. Then $\boldsymbol{X}$ and $\boldsymbol{Y}$ are almost surely asymptotically free.*

Chapter 5

System Design in the Presence of Phase Noise

Most certainly, one way of addressing the phase noise problem is to carefully design oscillators such that they have low and controlled levels of random phase variations. Such oscillators, in turn, can have higher power consumption, and can be costly. Given the ubiquity of wireless devices, and the exponential growth in their use, oscillator design has to be optimized in terms of cost and power. This renders the use of noisy oscillators inevitable. Therefore, it becomes important to appropriately design transceiver algorithms, and compensate for the effects of phase noise. In the sequel, we review prior work on designing systems in order to handle phase noise.

5.1 Design Approaches for SISO Systems with Phase Noise

The problem of designing wireless communication systems in the presence of phase noise has been investigated for decades. The main design approaches to this problem can be summarized as follows.

[a] Design phase noise trackers that track or estimate the phase noise process in the received signals, and compensate for its effects, followed by coherent detection of the transmitted symbols [3, 58].

[b] Design joint phase-estimation data-detection algorithms for compensating phase noise and detecting data [59].

[c] Design constellations that are optimized for the phase noise channel [60].

[d] Design error correcting codes that incorporate the effect of phase noise [61].

5.1.1 Phase Noise Tracking

We will first briefly review some methods for phase noise tracking/estimation used in communication receivers by considering the following question: How can phase noise estimators be designed such that optimal error rate performance can be achieved?

Trackers are used to track or estimate the phase noise based on the received samples, which are obtained after matched filtering and sampling of $r(t)$. The phase noise estimate is then used to compensate

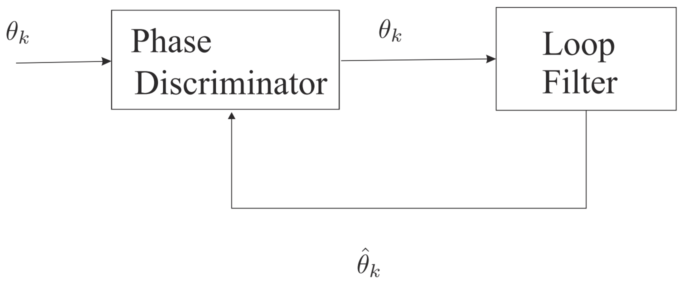


Figure 5.1: PLL Tracking

the received samples as

$$r_k \triangleq r_k e^{-j\hat{\theta}_k} = c_k e^{j\phi_k} + n_k \tag{5.1}$$

$$\phi_k \triangleq \theta_k - \hat{\theta}_k,\ n_k \triangleq \tilde{n_k} e^{-j\hat{\theta}_k}. \tag{5.2}$$

Here, $\hat{\theta}_k$ is the phase noise estimate, and ϕ_k is the residual phase error. Following this compensation, coherent detection of the transmitted symbols is performed by treating ϕ_k to be zero.

The most widely used tracker is the phase locked loop (PLL) [3, 58], which is shown in Fig. 5.1. Its operation can be summarized as follows: Let $\hat{\theta}_k$ be the tracked phase from a loop filter and θ_k be the phase noise in the received signal, which are the inputs to the phase discriminator. Let $\phi_k \triangleq \theta_k - \hat{\theta}_k$ denote the phase error process. This error signal is then fed to the loop filter, which produces $\hat{\theta}_k$, such that ϕ_k is minimized. When a PLL initially seeks to track the phase of the incoming signal, ϕ_k is large, which steadily decreases with time. This transient operating mode is called the acquisition mode of the PLL. When ϕ_k is small, the PLL is said to be locked to the incoming signal. Another tracker that is commonly used is the extended Kalman Filter (EKF) [50, 62], which has been shown to have a structure and performance similar to that of a PLL (refer to [63, 64]).

The performance of trackers can be evaluated by comparing their mean square error (MSE) with a lower bound on the phase noise estimation MSE. One way of characterizing the MSE lower bound is to evaluate the Bayesian Cramer-Rao bound (CRB) [65] for the phase noise model in (3.8). Particle filters [66], extended Kalman filters or smoothers, the MAP estimation algorithms in [64, 67] have been shown to achieve the CRB performance. Note that the closed-form analytical forms of the CRB are not available in general for the model in (3.8), when the data is unknown, or when the estimator has limited a priori information about the transmitted data [65].

In recent times, there has been significant efforts towards improving the performance of coded systems (like turbo codes) in the presence of phase noise. To address this problem, the per-survivor processing (PSP) algorithm proposed in [68] has been widely used. Here, phase noise estimation is first performed using an estimator like the PLL followed by sequence detection (using Viterbi or the BCJR algorithm) [69]. Another widely used technique for this problem is called turbo synchronization [70], where phase noise estimation is performed using the expectation-maximization (EM) algorithm. The phase estimates are then used to compute the a posteriori bit and symbol probabilities using algorithms like the BCJR [71–73]. In both PSP and turbo synchronization methods, the phase noise estimates obtained from the estimation algorithm are treated as the true value of phase noise. For MIMO systems with phase noise, similar design approaches have been reported in [28, 74, 75]. Note that the traditional approach can be

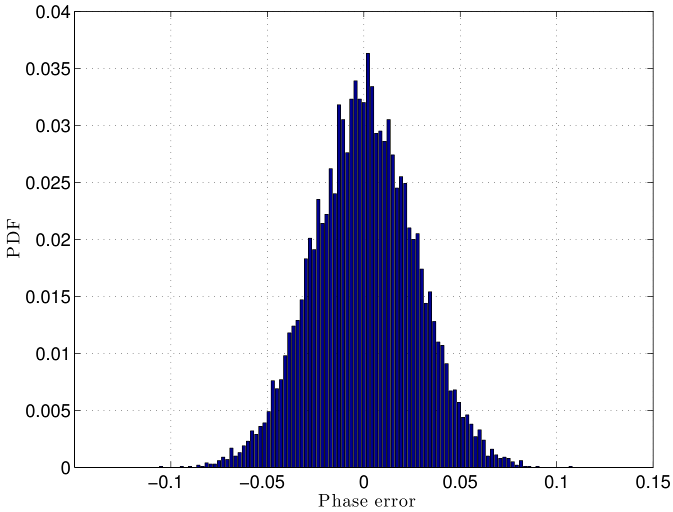


Figure 5.2: Pdf of phase error resulting from the compensation of the received signal with an EKF for $\sigma_\Delta^2 = 10^{-2}\text{rad}^2$.

viewed as a special case of the approach where algorithms for phase-estimation data-detection are jointly designed for compensating phase noise.

Phase Error Models

In the context of phase noise tracking, it is important to study models for the residual phase error process ϕ_k. So far, we considered receiver algorithms where ϕ_k is treated to be zero while performing coherent detection on the compensated received signals [71–73]. However, ϕ_k is an RV, and its statistics can be used for designing joint phase-estimation data-detection algorithms, which can significantly improve the error rate performance [60]. It is usually assumed that ϕ_k resulting from the estimator/tracker is Tikhonov distributed [76]. The Tikhonov or Von Mises pdf with circular mean 0 and variance $1/\rho$ is given as

$$p(\phi_k) = \frac{e^{\rho\cos(\phi_k)}}{2\pi I_0(\rho)},\ \phi_k \in [-\pi, \pi], \tag{5.3}$$

This pdf is approximately Gaussian for large values of ρ, and is also used to model the phase error after compensation using an estimator/tracker. Another pdf that is used to model ϕ_k is the wrapped Gaussian distribution [77],

$$p(\phi_k) = \frac{1}{\sqrt{2\pi\sigma_\text{p}^2}} \sum_{l\in\mathbb{Z}} e^{\frac{-(\phi_k - 2l\pi)^2}{2\sigma_\text{p}^2}},\ \phi_k \in [-\pi, \pi], \tag{5.4}$$

where σ_p^2 denotes the variance. As an example, we present the empirical phase error in Fig. 5.2, which is approximately Gaussian or Tikhonov distributed for a given symbol amplitude.

5.1.2 Joint Phase-Estimation Data-Detection Algorithms

We now review prior work, which addresses the following question: When the transmitted information signal is affected by AWGN and phase noise, how can a low-complexity joint phase noise estimation data-detection algorithms be designed such that (near) optimal system performance is achieved?

The problem of designing receiver algorithms that perform joint phase noise estimation and data detection in SISO links has been extensively studied, e.g., refer to [3, 58] and references therein. Some of the early works addressing this problem is [78, 79], which proposes simultaneous maximum-likelihood (ML) estimation of the data symbols, the carrier phase and the timing offset. In [80], MAP estimation based on the Viterbi algorithm is proposed for joint estimation of phase noise and data. The phase noise model considered in this work is similar to the random walk model in (3.8), but the innovations Δ_k are restricted to be discrete binary jumps. This shortcoming is addressed in [81], where the discrete Wiener process (3.8) is used. Specifically, the phase noise RV is assumed to be discrete in the range $[-\pi, \pi]$, and the Viterbi algorithm is used to determine the MAP phase noise and symbol estimates. A similar approach using the BCJR algorithm is proposed in [82]. The algorithms in [81, 82] are considered to be approximate implementations of the MAP symbol detector in this thesis, and are used as a benchmark when comparing the performance of various receiver algorithms. An analytical treatise of the MAP symbol detector for the system model in (3.9) can be found in [46], where it is shown that the optimal detector has a separable estimator-detector structure. The received signals are first used to compute the a posteriori pdf of phase noise. This phase noise pdf is then used for performing symbol detection. The problem of computing the a posteriori pdf of phase noise based on the received signals is shown to be intractable in general.

In order to derive the MAP symbol detector, it is possible to restrict the a posteriori phase noise pdf to a canonical family of distributions. This approach is reported in a much earlier work by Foschini *et al.* [60]. In their work, it is assumed that the phase of the received signal is tracked and compensated using a PLL. Then the a posteriori phase error pdf is approximated as a Tikhonov pdf [76], and is used to derive the ML detector. In a more recent effort, a similar detector is derived in [83] for the same phase noise model.

When the transmitted symbols are affected by random phase noise, methods based on the SPA [45] have been used for designing receiver algorithms. A joint phase-estimator data-detector based on the SPA, which is similar to an extended Kalman smoother is proposed in [62, 84]. In [59], the messages used in the SPA are restricted to be Tikhonov distributed. An extension of this approach is presented in [85] in order to handle both phase noise and a constant frequency offset. As a low complexity alternative to SPA, [50] employs the variational Bayesian framework. In [69], an algorithm based on the BCJR algorithm with forward and backward recursions is proposed for phase noise estimation and data detection. Applications of Monte Carlo sampling methods for joint phase noise estimation and data detection is investigated in [86] for both coded and uncoded systems.

Let us now see how the various low-complexity estimator-detectors proposed in prior work perform in terms of SEP with respect to the MAP algorithm [81, 82]. We consider uncoded data transmission of symbols from a 16-QAM constellation. The phase noise model used is the discrete Wiener phase noise model in (3.8) with $\sigma_\Delta^2 = 10^{-2}$ rad^2. The comparison is shown in Fig. 5.3, where we observe that the gap in performance between the various proposed algorithms and the MAP is significant. The gap in performance motivates the need to design new low-complexity algorithms for performing joint phase noise estimation and data detection for severely strong phase noise scenarios, particularly considering high order constellations. This is investigated in our works in Paper A [5, 48], which is appended to this thesis. Furthermore, a low-complexity phase noise estimator and data detector based on the SPA is proposed in Paper C for a MIMO system [6], which is impaired by phase noise as shown in Fig. 3.6(b).

5.1.3 Constellation Design

Another approach for improving system performance when affected by phase noise is to optimize the signal constellation that is used for transmission over the wireless link. In this regard, we summarize prior work that addresses the following question: How can two-dimensional signal constellations be

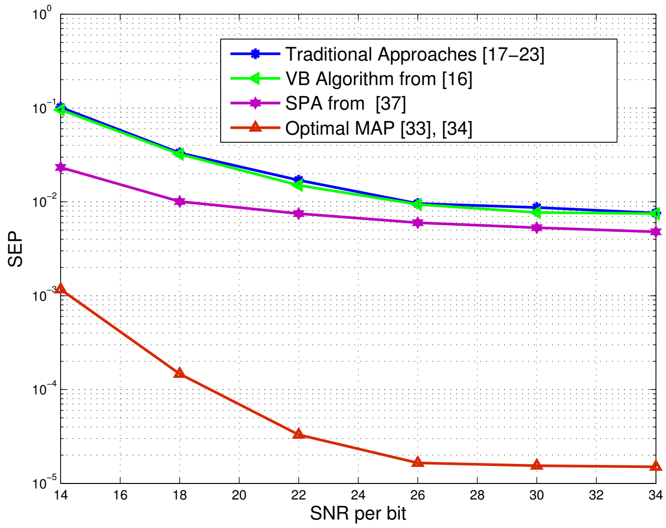


Figure 5.3: Comparison of the SEP performance between the various detectors and optimal MAP for 16-QAM, $\sigma_{\Delta}^2 = 10^{-2}\text{rad}^2$.

designed for channels with phase noise, such that a target objective function like error rate performance or the mutual information (MI) is optimized?

The problem of arranging M points in a two-dimensional plane such that a target objective function is optimized is a classical problem in communication theory [87]. For decades, this problem has been studied for different channel conditions and communication models [88–91]. SEP and mutual information (MI) are some of the performance measures that have been used as the target objective function. Constellations that minimize SEP for the phase noise channel are desirable in uncoded systems. Also, there are latency limited systems, and applications such as coordination of base stations in 4G cellular networks, and feedback loops in control systems that are preferably uncoded. In coded systems, some levels of processing such as clock recovery, forward error correcting (FEC) frame preamble decoding, and adaptive equalization depend on the SEP performance. Constellations that maximize the MI provide an upper bound on the achievable rate for any decoder [92], and is particularly relevant for symbol-based decoders such as in trellis-coded modulation or LDPC-based nonbinary coded schemes, and for systems that employ binary capacity-achieving codes like multilevel codes [93]. By properly designing non-binary codes to match the optimized constellations, or using binary multilevel codes, the MI of the constellation can be approached.

The design of constellations for wireless systems impaired by phase noise is addressed by Foschini *et al.* in [60]. In their work, an approximate MAP detector and its SEP are derived for the phase noise channel in (3.9). Then, constellations that minimize the SEP are obtained using a heuristic algorithm in [87]. In [94], constellations robust to phase noise are constructed heuristically such that they have low decoding complexity or simple decision regions (thus enabling quadrant or threshold-based decoding). In [95], the approximate SEP for a given phase offset in (3.9) is derived, which is minimized for designing constellations. In [96], a simple method for constructing spiral-shaped constellations is presented, and their performances are compared to that of the conventional constellations in the presence of memoryless phase noise. In a more recent effort [97], the problem of designing constellations that maximize the MI of the memoryless phase noise channel is addressed. In their work, first the (approximate) MI for the channel is derived, and constellations are optimized by maximizing the so derived MI using the simulated annealing algorithm.

Prior works have demonstrated that constellations designed for phase noise substantially outperform conventional constellations in terms of SEP and MI. However, in most prior work (except [60] and [97]) ad-hoc methods have been used. There has been very limited effort to address this problem based on rigorous optimization formulations. These factors motivate the need to revisit the problem of constellation design based on optimization formulations that use target objective functions like SEP or MI. This is the theme of our work in [7] that is appended to this thesis as Paper B. Using this approach, we demonstrate that the optimized constellations obtained outperform the conventional constellations, and those proposed in the literature.

5.1.4 Coding

Designing error correcting codes that achieve the channel capacity with arbitrarily small probability of error at affordable complexity is the holy grail for researchers in coding theory [98]. Though the problem of designing codes that are resilient to phase noise is not directly addressed in this thesis, it is relevant to shed light on prior work that has addressed the following question: How can capacity-achieving error correcting codes be designed for systems impaired by phase noise?

Designing error correcting codes such that they are amenable for phase noise scenarios is a challenging problem. In [83], the impact of phase noise on the error rate performance of standard error correcting codes is investigated. It is concluded that standard LDPC and turbo codes are effective in reducing the performance degradation incurred by phase noise. It is also noted that trellis coded modulation schemes experience significant performance degradation in the presence of strong phase noise. The design of codes for the phase noise channel is studied in [99–101]. These designs aided by phase noise estimation have been demonstrated to achieve good performance for channels with strong phase noise. In [102], LDPC codes are designed by dividing the codewords into sub-blocks such that the variation of phase noise over each sub-block is small. Phase estimates are then used to correct each sub-block, and phase ambiguity checks are applied using local check nodes. In [61], repeat-accumulate (RA) codes are designed where the phase ambiguity is resolved through differential encoding.

Chapter 6

Contributions and Future Directions

In this chapter, we summarize the papers that are appended to this thesis, and outline our main contributions.

[a] **Paper A: Soft metrics and their Performance Analysis for Optimal Data Detection in the Presence of Strong Oscillator Phase Noise**

In this paper, we address the classical problem of deriving an approximate MAP detector in the presence of random phase noise. We consider a system where the received signal that is impaired by phase noise is first compensated by a tracker, and then the resulting phase error pdf is assumed to be Gaussian distributed in order to derive an approximation of the MAP detector in [46]. Then, for the so derived detector, we analytically characterize the performance in terms of symbol error probability. Finally, by simulations, we show that the detector developed in our work outperforms those available in the literature for the considered signal constellations, phase noise scenarios and SNR values.

[b] **Paper B: Constellation Optimization in the Presence of Strong Phase Noise**

In this paper, we address the problem of optimizing signal constellations for strong phase noise. The problem is investigated by considering different optimization formulations, which provide an analytical framework for constellation design. The considered formulations optimize different objective functions such as the symbol error probability for the detector developed in Paper A, and the mutual information of the memoryless phase noise channel. We show that the optimized constellations significantly outperform conventional constellations and those proposed in the literature.

[c] **Paper C: Algorithms for Joint Phase Estimation and Decoding for MIMO Systems in the Presence of Phase Noise**

In this work, we derive the MAP symbol detector for the MIMO system in Fig. 3.6(b), where each transceiver antenna has its own oscillator. As in SISO systems [46], we observe that the computation of the optimum receiver is an analytically intractable problem, and is unimplementable in practice. In this light, we propose three suboptimal, low-complexity algorithms for approximately implementing the MAP symbol detector. These algorithms involve joint phase noise estimation and data detection. We obtain the first algorithm by means of the SPA, where we use the multivariate Tikhonov canonical distribution approach [59]. In our next algorithm, we derive an approximate MAP detector based on the smoother-detector framework developed in Paper A. The third algorithm is derived based on the variational Bayesian framework [50]. By simulations, we observe that the proposed techniques significantly outperform the other algorithms from prior works.

[d] **Paper D: Linear Precoding in the Presence of Phase Noise in Massive MIMO Systems: Design and Performance Analysis**

In this work, we study the impact of phase noise on the downlink performance of a MU-massive-MIMO system, where the base station employs a large number of transmit antennas M. We consider a setup as illustrated in Fig. 3.7, where the BS employs M_{osc} free-running oscillators, and M/M_{osc} antennas are connected to each oscillator. For this setup, we analyze the impact of phase noise on the performance of the zero-forcing (ZF), regularized ZF, and matched filter precoders when M and the number of users K are asymptotically large, while $M/K = \beta$ is a bounded constant. We analytically show that the impact of phase noise on the SINR can be quantified as an effective reduction in the quality of the CSI available at the BS when compared to a system without phase noise. The main result of this paper is that, for all the considered precoders, when β is small, the performance of the CO setup is superior to that of the DO setup. However, the opposite is true when β is large and the SNR at the users is low.

[e] **Paper E: On the Impact of Oscillator Phase Noise in Uplink Transmission for Massive MIMO-OFDM Systems**

In this work, we study the effect of oscillator phase noise on the uplink performance of a massive MIMO system, where a base station with a large number of antennas serves a single-antenna user. Specifically, we consider an OFDM-based uplink transmission, and analyze the effect of channel aging due to phase noise on the throughput performance for the CO and the DO setups. We derive the instantaneous SNR on each subcarrier, and analyze the ergodic capacity when a linear receiver is used. We discuss the averaging effects on phase noise in both setups, and finally, we propose a phase noise tracking algorithm based on Kalman filtering that mitigates the effect of channel aging due to phase noise on the system performance.

6.1 Future Work

Ongoing research and some possible topics for future research are described in the following:

- In the area of massive MIMO and phase noise, some of the relevant and interesting problems are:
  - Extend our work in Paper D to analyze the uplink performance of a MU-massive-MIMO system considering different oscillator setups at the base station.
  - Analyze the performance of hybrid precoders in a MU-massive-MIMO system in the presence of phase noise, with reduced RF chains and finite signal-conversion resolution.
  - Analyze the massive MIMO system performance considering amplifier nonlinearities and phase noise.
  - Develop blind, low-complexity phase noise estimation algorithms for massive MIMO systems.
  - Analyze the energy efficiency of a massive MIMO system considering imperfect oscillators (and the different setups), reduced RF chains, finite signal-converter resolution, and nonlinear amplifiers.
- In the context of MIMO systems, the following interesting problems have been identified:
  - Design differential space-time coding and detection methods in the presence of phase noise.
  - Design the MAP detector by accounting for both channel variations and phase noise.

  - Derive the approximate MAP detector using the SPA by considering an AR approximation for the phasor process.

- In the context of SISO systems, the following problems have been identified as challenging:
  - Design error correcting codes that incorporate the effect of phase noise impairments.
  - Design constellations for the Wiener phase noise channel.
  - Design approximate MAP detectors by considering multimodality in the phase error pdf or the a priori symbol probability distribution.
  - Design the MAP detector by accounting for both channel variations and phase noise.

References

[1] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, pp. 379–423, 1948.

[2] L. Hanzo, H. Haas, S. Imre, M. O'Brien, M. Rupp, and L. Gyongyosi, "Wireless myths, realities, and futures: From 3g/4g to optical and quantum wireless," *IEEE Proc. Special Centennial Issue*, vol. 100, pp. 1853–1888, May 2012.

[3] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*. Wiley, 1998.

[4] C. Liang, J. Jong, W. Stark, and J. East, "Nonlinear amplifier effects in communications systems," *IEEE Trans. Microwave Theory Tech.*, vol. 47, no. 8, pp. 1461–1466, Aug. 1999.

[5] R. Krishnan, M. R. Khanzadi, T. Eriksson, and T. Svensson, "Soft metrics and their performance analysis for optimal data detection in the presence of strong oscillator phase noise," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2385–2395, Jun. 2013.

[6] R. Krishnan, A. G. i Amat, T. Eriksson, and G. Colavolpe, "Algorithms for joint phase estimation and decoding for mimo systems in the presence of phase noise and quasi-static fading channels," *IEEE Trans. Signal Process., accepted for publication*, Jan. 2015.

[7] ——, "Constellation optimization in the presence of strong phase noise," *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 5056–5066, Dec. 2013.

[8] R. Krishnan, M. R. Khanzadi, N. Krishnan, Y. Wu, A. G. i Amat, T. Eriksson, and R. Schober, "Linear massive mimo precoders in the presence of phase noise – a large-scale analysis," *IEEE Trans. Veh. Tech., under review. [Online]. Available: http://arxiv.org/abs/1501.05461*, Jan. 2015.

[9] R. Krishnan, M. R. Khanzadi, N. Krishnan, A. G. i Amat, T. Eriksson, N. Mazzali, and G. Colavolpe, "On the impact of oscillator phase noise on the uplink performance in a massive mimo-ofdm system," *IEEE Signal Process. Lett., under review. [Online]. Available: http://arxiv.org/abs/1405.0669.*, May. 2014.

[10] A. Hajimiri and T. H. Leee, "A general theory of phase noise in electrical oscillators," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 2, pp. 179 – 194, Feb. 1998.

[11] G. V. Klimovitch, "A nonlinear theory of near-carrier phase noise in free-running oscillators," *Proceedings of the 2000 Third IEEE International Caracas Conference on Devices, Circuits and Systems*, pp. 1–6, 2000.

[12] A. Chorti and M. Brookes, "A spectral model for rf oscillators with power-law phase noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 9, pp. 1989–1999, Sep. 2006.

[13] J. Dauwels, "On graphical models for communications and machine learning: Algorithms, bounds, and analog implementation," *Ph.D. dissertation 16365*, vol. ETH Zurich, Switzerland, 2005.

[14] A. Demir, "Computing timing jitter from phase noise spectra for oscillators and phase-locked loops with white and 1/f noise," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 53, no. 9, pp. 1869 – 1884, Sep. 2006.

[15] D. Lee, "Analysis of jitter in phase-locked loops," *IEEE Trans. Circuits Syst. II*, vol. 49, no. 11, pp. 704 – 711, Nov. 2002.

[16] M. R. Khanzadi, A. Panahi, D. Kuylenstierna, and T. Eriksson, "A model-based analysis of phase jitter in rf oscillators," *IEEE International Frequency Control Symposium (FCS), 2012*, pp. 1–4, May 2012.

[17] R. Poore, "Phase noise and jitter," *Agilent EEsof EDA*, 2001.

[18] F. Rusek, D. Persson, K. L. Buon, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up mimo: opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[19] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[20] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.

[21] H. Ghozlan and G. Kramer, "On Wiener phase noise channels at high signal-to-noise ratio," *Global Telecom. Conf. (GLOBECOM)*, p. 19191924, Dec. 2013.

[22] A. Lapidoth, *A Foundation in Digital Communication*, 1st ed. Cambridge Univ. Press, 2009.

[23] D. S. Baum and H. Bolcskei, "Impact of phase noise on MIMO channel measurement accuracy," in *Vehicular Tech. Conf. (VTC)*, vol. 3, Sep. 2004, pp. 1614 – 1618.

[24] W. Clarke, "A statistical theory of mobile-radio reception," *Bell Syst. Tech. J.*, vol. 47, no. 6, pp. 957–1000, Jul./Aug. 1968.

[25] F. Simoens and M. Moeneclaey, "Code-aided estimation and detection on time-varying correlated mimo channels: A factor graph approach," *EURASIP J. Applied Signal Processing*, no. 1, pp. 1–11, 2006.

[26] F. Bohagen, P. Orten, and G. E. Oien, "Design of capacity-optimal high-rank line-of-sight mimo channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 790–804, Apr. 2007.

[27] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive mimo: the next wireless revolution?" *IEEE Commun. Mag.*, vol. 52, no. 9, pp. 56–62, Sep. 2014.

[28] H. Mehrpouyan, A. A. Nasir, S. D. Blostein, T. Eriksson, G. Karagiannidis, and T. Svensson, "Joint estimation of channel and oscillator phase noise in mimo systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sep. 2012.

[29] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[30] R. Krishnan, M. R. Khanzadi, L. Svensson, T. Eriksson, and T. Svensson, "Variational bayesian framework for receiver design in the presence of phase noise in mimo systems," *IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 347–352, Apr. 2012.

[31] K. Huber and S. Haykin, "Improved bayesian mimo channel tracking for wireless communications: incorporating a dynamical model," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2458–2466, Sep. 2006.

[32] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.

[33] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, “Multiuser mimo achievable rates with downlink training and channel state feedback,” *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[34] E. Bjornson, J. Hoydis, M. Kountouris, and M. Debbah, “Massive mimo systems with non-ideal hardware: energy efficiency, estimation, and capacity limits,” *IEEE Trans. Inf. Theory, Under Review, [Online]. Available: http://arxiv.org/abs/1307.2584.*, Jul. 2013.

[35] A. Pitarokoilis, S. K. Mohammed, and E. G. Larsson, “Uplink performance of time-reversal mrc in massive mimo system subject to phase noise,” *IEEE Trans. Wireless Commun., accepted for publication. [Online]. Available: http://arxiv.org/abs/1306.4495.*, 2014.

[36] K. T. Truong and R. W. H. Jr., “Effects of channel aging in massive mimo systems,” *IEEE J. Commun. Netw. (Special Issue on Massive MIMO)*, vol. 15, no. 4, pp. 338–351, Aug. 2013.

[37] T. C. W. Schenk, X.-J. Tao, P. F. M. Smulders, and E. R. Fledderus, “On the influence of phase noise induced ici in mimo ofdm systems,” *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 682–684, Aug. 2005.

[38] T. Tao, *Topics in random matrix theory*. Graduate Studies in Mathematics, AMS, 2012, vol. 132.

[39] B. D. O. Anderson and J. B. Moore, *Optimal filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1979.

[40] P. Maybeck, *Stochastic models, estimation and control*. Academic Press, 1982, vol. 3.

[41] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking,” *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–189, Feb. 2002.

[42] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer-Verlag, 1994.

[43] H. Urkowitz, “Energy detection of unknown deterministic signals,” in *Proc. IEEE*, vol. 55, no. 4, Apr. 1967, p. 523531.

[44] M. K. Simon, F. F. Digham, and M. Alouini, “On the energy detection of unknown signals over fading channels,” *ICC 2003 Conf. Record, Anchorage, Alaska*, pp. 11–15, May 2003.

[45] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[46] P. Y. Kam, S. S. Ng, and T. S. Ng, “Optimum symbol-by-symbol detection of uncoded digital data over the gaussian channel with unknown carrier phase,” *IEEE Trans. Commun.*, vol. 42, no. 8, pp. 2543–2552, Aug. 1994.

[47] M. J. Beal, “Variational algorithms for approximate bayesian inference,” *PhD. dissertation, Univ. Coll. London*, 1998.

[48] R. Krishnan, H. Mehrpouyan, T. Eriksson, and T. Svensson, “Optimal and approximate methods for detection of uncoded data with carrier phase noise,” *IEEE Global Telecomm. Conf. (GLOBECOM)*, pp. 1–6, Houston, 5-9 Dec. 2011.

[49] P. Worthen and W. E. Stark, “Unified design of iterative receivers using factor graphs,” *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 843–849, Feb. 2001.

[50] M. Nissila and S. Pasupathy, “Adaptive iterative detectors for phase-uncertain channels via variational bounding,” *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 716–725, Mar. 2009.

[51] M. Katz and S. Shamai, “On the capacity-achieving distribution of the discrete-time noncoherent and partially coherent awgn channels,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2257 – 2270, Oct. 2004.

[52] P. Hou, B. Belzer, and T. R. Fischer, “Shaping gain of the partially coherent additive white gaussian noise channel,” *IEEE Commun. Lett.*, vol. 6, pp. 175–177, 2002.

[53] A. Lapidoth, “On phase noise channels at high snr,” *IEEE Proc. Information Theory Workshop*, Bangalore, India, Oct. 2002.

[54] M. R. Khanzadi, G. Durisi, and T. Eriksson, “Capacity of multiple-antenna phase-noise channels with common/separate oscillators,” *IEEE Trans. Commun., submitted*, vol. [Online]. Available: http://arxiv.org/abs/1409.0561., Sep. 2014.

[55] G. Akemann, J. Baik, and E. P. Di Francesco, *The Oxford handbook of random matrix theory*. Oxford University Press, 2011.

[56] A. M. Tulino and S. Verdu, *Random matrix theory and wireless communications*, 1st ed. Delft, Netherlands: Now Publishers Inc., 2004.

[57] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, “Large system analysis of linear precoding in correlated miso broadcast channels under limited feedback,” *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[58] U. . Mengali and A. N. Andra, *Synchronization Techniques for Digital Receivers*, 1st ed. New York: Plenum Press, 1997.

[59] G. Colavolpe, A. Barbieri, and G. Caire, “Algorithms for iterative decoding in the presence of strong phase noise,” *IEEE Journ. Sel. Areas Commun.*, vol. 23, no. 9, pp. 1748–1757, Sep. 2005.

[60] G. J. Foschini, R. D. Giltlin, and S. B. Weinstein, “On the selection of a two-dimensional signal constellation in the presence of phase jitter and gaussian noise,” *Bell Syst. Tech. J.*, vol. 52, pp. 927 – 965, Jul/Aug. 1973.

[61] A. Barbieri and G. Colavolpe, “On the information rate and repeat-accumulate code design for phase noise channels,” *IEEE Trans. Commun.*, vol. 59, no. 12, pp. 3223 – 3228, Dec. 2011.

[62] J. Dauwels and H. A. Loeliger, “Joint decoding and phase estimation: An exercise in factor graphs,” *Proc. IEEE Symp. Inf. Theory*, p. 231, Jun. 2003.

[63] A. Patapoutian, “On phase-locked loops and kalman filter,” *IEEE Trans. Commun.*, vol. 47, no. 5, pp. 670–672, May 1999.

[64] R. Bucy and H. Youssef, “Optimal phase demodulation,” *IEEE Trans. Auto. Cont.*, vol. 21, no. 5, pp. 732–737, Oct. 1976.

[65] S. Bay, C. Herzet, J. M. Brossier, J. P. Barbot, and B. Geller, “Analytic and asymptotic analysis of bayesian cramer rao bound for dynamical phase offset estimation,” *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 61–70, Sep. 2008.

[66] P. Amblard, J. Brossier, and E. Moisan, “Phase tracking: What do we gain from optimality? particle filtering versus phase-locked loops,” *Signal Process.*, vol. 83, pp. 151–167, Oct. 2003.

[67] J. Bhatti and M. Moeneclaey, “Feedforward data-aided phase noise estimation from a dct basis expansion,” *EURASIP J. Wireless Commun. Net., accepted for publication. [Online]. Available: http://arxiv.org/abs/1306.4495.*, vol. 2009, pp. 1–11, Jan. 2009.

[68] R. Raheli, A. Polydoros, , and C.-K. Tzou, “Per-survivor processing: a general approach to mlse in uncertain environments,” *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 354–364, Apr. 1995.

[69] A. Anastasopoulos and K. M. Chugg, “Adaptive iterative detection for phase tracking in turbo-coded systems,” *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2135–2144, Dec. 2001.

[70] C. H. *et al.*, “Code-aided turbo synchronization,” *IEEE Proceed.*, vol. 2007, pp. 1255–1271, Jun. 2007.

[71] V. Lottici and M. Luise, “Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations,” *IEEE Trans. Commun.*, vol. 52, no. 4, pp. 661–669, Apr. 2004.

[72] N. N. *et al.*, “A theoretical framework for soft information based synchronization in iterative (turbo) receivers,” *EURASIP J. Wireless Commun. Netw.*, vol. 2005, pp. 117–127, Apr. 2005.

[73] G. Ferrari, G. Colavolpe, and R. Raheli, “On linear predictive detection for communications with phase noise and frequency offset,” *IEEE Trans. Veh. Techol.*, vol. 56, no. 4, pp. 2073 – 2085, Jul. 2007.

[74] A. A. Nasir, H. Mehrpouyan, R. Schober, and Y. Hua, “Phase noise in mimo systems: Bayesian cramer rao bounds and soft-input estimation,” *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.

[75] N. Hadaschik, O. H. U. K. G. A. M. Dorpinghaus, A. Senst, and H. Meyr, “Phase noise in mimo systems: Bayesian cramer rao bounds and soft-input estimation,” *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2675–2692, May 2013.

[76] A. J. Viterbi, “Phase-locked loop dynamics in the presence of noise by fokker-planck techniques,” *Proceed. IEEE*, vol. 51, no. 12, pp. 1737–1753, Dec. 1963.

[77] B. G. *et al.*, “Calculation of mutual information for partially coherent gaussian channels with application to fiber optics,” *IEEE Trans. Inf. Theory*, vol. 57, no. 9, pp. 5720–5736, Sep. 2011.

[78] H. Kobayashi, “Simultaneous adaptive estimation and decision algorithm for carrier modulated data transmission systems,” *IEEE Trans. Commun. Tech.*, vol. 19, no. 3, pp. 268–280, Jun. 1971.

[79] D. Falconer and J. Salz, “Optimal reception of digital data over the gaussian channel with unknown delay and phase jitter,” *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 117–126, Jan. 1977.

[80] G. Ungerboeck, “New application for the viterbi algorithm: Carrier phase tracking in synchronous data transmission systems,” *Proc. Nat. Telecomm. Conf.*, pp. 734 – 738, Apr. 1974.

[81] O. Macchi and L. L. Scharf, “A dynamic programming algorithm for phase estimation and data decoding on random phase channels,” *IEEE Trans. Inf. Theory*, vol. 27, no. 5, pp. 581–595, Sep. 1981.

[82] M. Peleg, S. S. (Shitz), and S. Galn, “Iterative decoding for coded noncoherent mpsk communications over phase-noisy awgn channel,” *IEE Proc. Commun.*, vol. 147, pp. 87 – 95, Apr. 2000.

[83] T. Minowa, H. Ochiai, and H. Imai, “Phase-noise effects on turbo trellis-coded modulation over m-ary coherent channels,” *IEEE Trans. Commun.*, vol. 52, no. 8, Aug. 2004.

[84] J. Dauwels and H. A. Loeliger, “Phase estimation by message passing,” *Proc. IEEE Int. Conf. Commun, (ICC ’04)*, pp. 523 – 527, 2004.

[85] G. Colavolpe, A. Barbieri, and G. Caire, "Joint iterative detection and decoding in the presence of phase noise and frequency offset," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 171 – 179, Jan. 2007.

[86] F. S. *et al.*, "Monte carlo solutions for blind phase noise estimation," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–11, Jan. 2009.

[87] B. W. Kernighan and S. Lin, "Heuristic solution of a signal design optimization problem," *Bell Syst. Tech. J.*, vol. 52, no. 7, pp. 1145 – 1159, Jun. 1973.

[88] G. F. Foschini, R. D. Gitlin, and S. B. Weinstein, "Optimization of two dimensional signal constellations in the presence of gaussian noise," *Bell Syst. Tech. J.*, vol. 22, no. 1, pp. 28–38, Jan. 1974.

[89] L. Beygi, E. Agrell, and M. Karlsson, "Optimization of 16 - point ring constellations in the presence of nonlinear phase noise," *Proc. Opt. Fib. Comm. Conf. (OFC)*, Mar. 2011.

[90] T. Pfau, X. Liu, and S. Chandrasekhar, "Optimization of 16 - ary quadrature amplitude modulation constellations for phase noise impaired channels," *ECOC 11, Optical Society of America, Tech. Dig.*, Tu.3.A.6 2011.

[91] C. Haeger, A. G. i Amat, A. Alvarado, and E. Agrell, "Constellation optimization for coherent optical channels distorted by nonlinear phase noise," *in IEEE Proc. Global Commun. Conf.*, Anaheim, CA, Dec. 2011.

[92] M. F. Barsoum, C. Jones, and M. Fitz, "Constellation design via capacity maximization," *IEEE Int. Sym. Inf. Th. (ISIT)*, 2007.

[93] U. Wachsmann, R. Fischer, and J. B. Huber, "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361–1391, Jul. 1999.

[94] S. Hulyalkar, "64 qam signal constellation which is robust in the presence of phase noise and has decoding complexity," *U.S.Patent*, no. 5832041, Nov. 1998.

[95] Y. Li, S. Xu, and H. Yang, "Design of circular signal constellations in the presence of phase noise," *4th Int. Conf. Wireless Commun., Networking, and Mobile Computing, IEEE WiCOM'08*, pp. 1–8, 2008.

[96] B. Kwak, N. Song, B. Park, and D. S. Kwon, "Spiral qam: A novel modulation scheme robust in the presence of phase noise," *IEEE 68th Veh. Tech. Conf. 2008, VTC 2008-Fall*, pp. 1–5, 2008.

[97] F. Kayhan and G. Montorsi, "Constellation design for memoryless phase noise channels," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2874–2883, May 2014.

[98] S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2004.

[99] M. Ferrari and A. Tomasoni, "Low-complexity z_4 ldpc code design under a gaussian approximation," *IEEE Wireless Comm. Lett.*, no. 99, pp. 1–4, 2012.

[100] A. Barbieri and G. Colavolpe, "Soft-output decoding of rotationally invariant codes over channels with phase noise," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2125–2133, Nov. 2007.

[101] Y. Aviv and I. Rosenhouse, "Bit-error-rate of ldpc coded qam in the presence of a residual phase noise process and a non-linear distortion," *IEEE 25th Convention of Electrical and Electronics Engineers in Israel*, pp. 175–179, Dec. 2008.

[102] S. Karuppasami and W. Cowley, "Construction and iterative decoding of ldpc codes over rings for phase-noisy channels," *EURASIP J. Wireless Commun.*, vol. 2008, pp. 1–9, 2008.