



CHALMERS

Chalmers Publication Library

Variational Bayesian Framework for Receiver Design in the Presence of Phase Noise in MIMO Systems

This document has been downloaded from Chalmers Publication Library (CPL). It is the author's version of a work that was accepted for publication in:

IEEE Wireless Communications and Networking Conference, WCNC 2012, Paris, 1-4 April 2012 (ISSN: 1525-3511)

Citation for the published paper:

Krishnan, R. ; Khanzadi, M. ; Svensson, L. (2012) "Variational Bayesian Framework for Receiver Design in the Presence of Phase Noise in MIMO Systems". IEEE Wireless Communications and Networking Conference, WCNC 2012, Paris, 1-4 April 2012 pp. 347-352.

<http://dx.doi.org/10.1109/WCNC.2012.6214388>

Downloaded from: <http://publications.lib.chalmers.se/publication/153458>

Notice: Changes introduced as a result of publishing processes such as copy-editing and formatting may not be reflected in this document. For a definitive version of this work, please refer to the published source. Please note that access to the published version might require a subscription.

Chalmers Publication Library (CPL) offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all types of publications: articles, dissertations, licentiate theses, masters theses, conference papers, reports etc. Since 2006 it is the official tool for Chalmers official publication statistics. To ensure that Chalmers research results are disseminated as widely as possible, an Open Access Policy has been adopted. The CPL service is administrated and maintained by Chalmers Library.

(article starts on next page)

Variational Bayesian Framework for Receiver Design in the Presence of Phase Noise in MIMO Systems

Rajet Krishnan, M. Reza Khanzadi, Lennart Svensson, Thomas Eriksson, and Tommy Svensson

Department of Signals and Systems, Chalmers University of Technology

E-mail: {raj, khazadi, lennart.svensson, thomase, tommy.svensson}@chalmers.se

Abstract—In this work, the problem of receiver design for phase noise estimation and data detection in the presence of oscillator phase noise in a point-to-point *multiple-input multiple-output* (MIMO) system is addressed. First, we discuss some interesting and challenging aspects in receiver design for MIMO systems in the presence of Wiener phase noise. Then, using the *variational Bayesian* (VB) framework, a joint iterative phase noise estimator and symbol detector are developed based on inverse Gibbs or variational free energy maximization. Further, the *symbol error probability* (SEP) of the newly proposed iterative scheme is compared with the optimal *maximum likelihood* (ML) detector with perfect phase information for 16-phase shift keying (PSK) and 16-quadrature amplitude modulation (QAM) schemes.

I. INTRODUCTION

Employing *multiple-input multiple-output* (MIMO) system for wireless communication has proven to significantly enhance performance in terms of capacity and link reliability in fading environments [1]. However, performance analysis of MIMO systems is based on the assumption that carrier phase is known to the receiver, and that the *random, time varying* phase difference between the transmitter and receiver are essentially absent. Practical receiver designs for MIMO systems based on this assumption can result in significant performance loss, and have to be appropriately addressed [2].

The problem of receiver design in the presence of random, time varying phase noise originating from the local oscillator has been studied extensively for *single-input single-output* (SISO) systems. We refer the readers to [2]–[5] and the references therein for relevant work. Receiver designs in the presence of random phase noise have also received considerable attention in *orthogonal frequency-division multiplexing* (OFDM) [12] and MIMO-OFDM [6] systems. Though some authors have addressed various issues concerning phase noise in MIMO systems [7], [8], there is severely limited understanding in literature about receiver design for MIMO systems in the presence of time varying random phase noise. The work in [7] discusses the impact of random phase noise on MIMO channel measurement systems and the estimated capacity. The impact of phase noise from free running oscillator in MIMO systems on estimated channel capacity, doppler of multipath components, and delay is discussed in [8].

In general, much of the research focus for receiver design in a MIMO setting has been on developing receiver algorithms for *joint channel estimation and data detection*, refer to [9] and the references therein. In this context, there exists a perception that the problem of receiver design in MIMO in the presence of random phase noise (from the oscillator), is effectively a channel estimation/data detection problem, since the impact of phase noise can be moved into the channel matrix [7].

In this work, we consider the problem of receiver design in the presence of *random, time varying* phase noise arising from oscillator instabilities in a point-to-point MIMO system. In the

ensuing sections, we first motivate the problem in a MIMO setting and discuss why the problem is interesting and different from the channel estimation/data detection problem in some cases of interest. Specifically we discuss scenarios where the impact of phase noise need not be moved into the channel matrix that in turn results in significant complexity reduction of the problem. Further we develop an iterative phase noise estimation and data detection scheme based on the *Variational Bayesian* (VB) framework [10], [11] and investigate its performance in the presence of Rayleigh fading channels. The VB theory provides an optimization framework to approximate inference algorithms involving computation of marginal *probability distribution functions* (pdf) [11]. The framework has been used for developing highly efficient receiver algorithms with parameter uncertainty for SISO and OFDM systems in [4], [12] respectively.

The contributions and organization of this paper are as follows - In Section II, the MIMO system model under study is presented. Based on the system model, we present simple illustrations in Section III to demonstrate the new challenges in receiver design in the presence of random time varying phase noise in MIMO. Further, in sections IV and V, we discuss the VB framework and analytically derive the structure of the receiver comprised of an iterative phase estimator and symbol detector. In Section VI, we present our simulation results to demonstrate the performance of the proposed receiver and give interesting insights into the problem.

II. SYSTEM MODEL

Consider a point-to-point MIMO system with N_t transmit antennas and N_r receive antennas. Data is transmitted from each transmit antenna in the form of frames, each of length N_f , and is uncoded. Further, the symbols transmitted from each transmit antenna are independent of each other. The channel between the transmit and receiver antennas are assumed to be known (estimated). As we shall see in the sequel, this is a realistic assumption and corresponds to the scenario where the channel varies much slower than the phase noise process, and has been already estimated. Each of the transmit and receive antennas is equipped with an independent free-running oscillator, perturbed by random phase noise process modeled as a random walk. The received signal model at the k th antenna at time instant n is modeled as follows

$$\begin{aligned} y_k(n) &= \sum_{l=1}^{N_t} e^{j\theta_k^{[r]}(n)} h_{lk}(n) e^{j\theta_l^{[t]}(n)} s_l(n) + w_k(n), \quad (1) \\ &= \sum_{l=1}^{N_t} h_{lk}(n) e^{j\theta_{lk}(n)} s_l(n) + w_k(n), \\ l &= 1, \dots, N_t, k = 1, \dots, N_r, n = 1, \dots, N_f, \end{aligned}$$

where $\theta_{lk}(n)$ is the sum of transmit and receive phase noise samples at the l th transmit antenna and k th receive antenna at time instant n , and is also referred to as the phase noise in each link. The phase noise samples at the l th transmit antenna and the k th receive antenna at time instant n are denoted as $\theta_l^{[t]}(n)$ and $\theta_k^{[r]}(n)$ respectively. Both $\theta_l^{[t]}(n), \forall l \in \{1, \dots, N_t\}$ and $\theta_k^{[r]}(n), \forall k \in \{1, \dots, N_r\}$ are assumed to be samples from a Wiener phase noise process given as

$$\begin{aligned}\theta_l^{[t]}(n) &= \theta_l^{[t]}(n-1) + \Delta_{l,t}(n), \\ \theta_k^{[r]}(n) &= \theta_k^{[r]}(n-1) + \Delta_{k,r}(n),\end{aligned}$$

where $\Delta_{l,t}(n) \sim \mathcal{N}(0, \sigma_{\Delta_{l,t}}^2)$ and $\Delta_{k,r}(n) \sim \mathcal{N}(0, \sigma_{\Delta_{k,r}}^2)$. Further, the phase noise samples in all the antennas are assumed to be varying from symbol-to-symbol, but constant over a symbol period. The data transmitted from the l th transmit antenna at time instant n is uncoded and drawn from an M -ary signal constellation with equal probability, and denoted as $s_l(n)$; i.e., $s_l(n) \in \mathcal{C}$, where \mathcal{C} is the set of all symbols in the signal constellation. The channel gain from the l th transmit antenna to the k th receive antenna is denoted as $h_{lk}(n) \sim \mathcal{CN}(0, \sigma_{h_{lk}}^2)$. $w_k(n) \sim \mathcal{CN}(0, \sigma_w^2)$ is the zero-mean *additive white Gaussian noise* (AWGN) at the k th receive antenna at time instant n . At time instant n , we denote the vector of all received signals at all receive antennas as $\mathbf{y}(n) = [y_1(n), \dots, y_{N_r}(n)]$, vector of phase noise samples as $\boldsymbol{\theta}(n) = \{\theta_{lk}(n)\}$, and the transmitted symbol vector as $\mathbf{s}(n) = [s_1(n), \dots, s_{N_t}(n)], l = 1, \dots, N_t, k = 1, \dots, N_r, n = 1, \dots, N_f$.

III. RECEIVER DESIGN IN MIMO - CHALLENGES

Based on the MIMO system model above, we now present some interesting aspects and challenges in receiver design for MIMO systems in the presence of Wiener phase noise. It is generally perceived that the problem of receiver design in the presence of random phase noise can be handled by moving the impact of phase noise (from the oscillator) into the channel matrix [7]. Then the only novel aspect of the problem of receiver design in the presence of phase noise would be to include the characteristics of the Wiener process in an equivalent channel estimation/data detection problem. In the sequel, using a few illustrations, we demonstrate that in some scenarios, the receiver design problem in the presence of time varying phase noise process is different, and can be handled differently from a channel estimation/data detection problem.

- There are several scenarios of interest, where the Wiener phase noise process drifts much faster than a random channel process. For the receiver design problem involving phase estimation and data detection in these scenarios, the channel can be assumed to be known (estimated) and static, and the phase noise impact need not be moved into the channel matrix. Consider, e.g., a system with point-to-point communication, transmitting at a symbol rate of 10^6 symbols/sec. Let the center carrier frequency for transmission be 2.4 GHz and the relative velocity between the communicating point be 5 km/hr. Then the maximum doppler spread for this link is given as $f_m \approx 13$ Hz. Let the sampling interval be $T_s = 10^{-6}$ s, variance of the innovation of phase noise process be $\sigma_{\Delta_{l,t}}^2 = \sigma_{\Delta_{k,r}}^2 = 1 \times 10^{-3}$. Then the 3dB bandwidth of the phase noise process is $f_{3db} = \frac{\sigma_{\Delta_{k,r}}^2}{2\pi T_s} \approx 160$ Hz [13]. Hence for small relative velocity, of the range of pedestrian velocity, the phase noise process can be assumed to vary

much faster than the channel. This is because in these scenarios, the bandwidth of the Wiener phase noise process is much higher than the doppler spread.

- For the case of channels (known and static) as discussed, the actual number of phase variables to be estimated can be reduced to $N_t + N_r$, as opposed to estimating $N_t \times N_r$ variables. This is because though the number of phases to be estimated is $N_t \times N_r$, only $N_t + N_r$ of the variables are relatively more fast varying. Hence by excluding them from the channel matrix, the problem is made considerably less complex.¹
- Oscillator phase noise in MIMO systems result in *both phase distortions and amplitude distortions*, unlike in the case of SISO systems. This can be illustrated using a simple example :- for a 2×2 MIMO system, the received signal at receive antenna $k = 1$ at high SNR (without the time index n) can be written as

$$\begin{aligned}y_1 &\approx e^{j\theta_{11}} h_{11} s_1 + e^{j\theta_{21}} h_{21} s_2, \\ y_1 y_1^* &= s_1 s_1^* h_{11} h_{11}^* + s_2 s_2^* h_{21} h_{21}^* \\ &\quad + 2\Re\{s_1 s_2^* h_{11} h_{21}^* e^{j(\theta_{11} - \theta_{21})}\}.\end{aligned}\quad (2)$$

As evident from (2), the amplitude of the received signal depends on the phase difference between the signals arriving at the receive antenna. This is irrespective of whether an equal energy constellation like PSK or a non-equal energy constellation like QAM is used.

IV. VARIATIONAL BAYESIAN FRAMEWORK - THEORY

An optimal *maximum a-posteriori* (MAP) symbol detector for transmitted from all transmit antennas at time instant n is given as

$$\begin{aligned}\hat{\mathbf{s}}(n) &= \underset{\mathbf{s}(n) \in \mathcal{C}}{\operatorname{argmax}} P(\mathbf{s}(n) | \mathbf{y}(n)), \\ &= \underset{\mathbf{s}(n) \in \mathcal{C}}{\operatorname{argmax}} \int_{\boldsymbol{\theta}(n)} P(\mathbf{s}(n), \boldsymbol{\theta}(n) | \mathbf{y}(n)) d\boldsymbol{\theta}(n), \\ &= \underset{\mathbf{s}(n) \in \mathcal{C}}{\operatorname{argmax}} \int_{\boldsymbol{\theta}(n)} P(\mathbf{s}(n) | \boldsymbol{\theta}(n), \mathbf{y}(n)) P(\boldsymbol{\theta}(n) | \mathbf{y}(n)) d\boldsymbol{\theta}(n).\end{aligned}\quad (3)$$

The detector in (3) is analytically intractable, and to work around this intractability, approximate probabilistic graphical models and inference techniques are widely used. The variational Bayesian approach [10] is one such probabilistic inference technique that is used for approximate inference and learning. This technique provides an optimization-based framework for approximating marginal pdfs and have been used to develop effective detectors for SISO and OFDM systems [4], [12] for time varying phase noise. The technique can be explained as follows - consider the problem of computing the log likelihood of $\mathbf{y}(n)$. Dropping the

¹ It is possible to achieve further reduction to $N_t + N_r - 1$ variables. This is by subtracting all transmit phases by any one of the transmit phases and adding the same amount to all the receive antenna phases. For instance, in (1) the phase states to be estimated are $\{\theta_{lk}\}$, $l = 1, \dots, N_t, k = 1, \dots, N_r$ implying that there are $N_t \times N_r$ phase noise variables to be estimated (in $N_t \times N_r$ links). However, the transmit phase states can be transformed to $\{0, \theta_l^{[t]} - \theta_1^{[t]}\}$, $l = 2, \dots, N_t$, and the receive phase states can be changed to $\{\theta_k^{[r]} + \theta_1^{[t]}\}$, $k = 1, \dots, N_r$. This transformation in effect produces the same received signal model in (1), though the transmit and receive phase states to be estimated have been altered and reduced to $N_t + N_r - 1$ states.

time index n for convenience and applying Jensen's inequality, the log likelihood can be lower bounded as follows

$$\begin{aligned} \log P(\mathbf{y}) &= \log \sum_{\mathbf{s}} \int_{\boldsymbol{\theta}} Q(\mathbf{s}, \boldsymbol{\theta}) \frac{P(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y})}{Q(\mathbf{s}, \boldsymbol{\theta})} d\boldsymbol{\theta}, \\ &\geq \sum_{\mathbf{s}} \int_{\boldsymbol{\theta}} Q(\mathbf{s}, \boldsymbol{\theta}) \log \frac{P(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y})}{Q(\mathbf{s}, \boldsymbol{\theta})} d\boldsymbol{\theta}. \end{aligned} \quad (4)$$

When the variational distribution $Q(\mathbf{s}, \boldsymbol{\theta})$ is set as $P(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y})$, the lower bound in (4) is achieved. However, the algorithm is restricted to search over a family of factorized distributions of the form: $Q(\mathbf{s}, \boldsymbol{\theta}) = q_{\mathbf{s}}(\mathbf{s})q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. This corresponds to assuming that \mathbf{s} and $\boldsymbol{\theta}$ being independent of each other given the observation \mathbf{y} . Hence the lower bound is given by

$$\begin{aligned} \log P(\mathbf{y}) &\geq \sum_{\mathbf{s}} \int_{\boldsymbol{\theta}} q_{\mathbf{s}}(\mathbf{s})q_{\boldsymbol{\theta}}(\boldsymbol{\theta}) \log \frac{P(\mathbf{s}, \boldsymbol{\theta}, \mathbf{y})}{q_{\mathbf{s}}(\mathbf{s})q_{\boldsymbol{\theta}}(\boldsymbol{\theta})} d\boldsymbol{\theta}, \\ &\triangleq \mathcal{H}(q_{\mathbf{s}}(\mathbf{s}), q_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \mathbf{y}). \end{aligned} \quad (5)$$

Here, $\mathcal{H}(q_{\mathbf{s}}(\mathbf{s}), q_{\boldsymbol{\theta}}(\boldsymbol{\theta}), \mathbf{y})$ is referred to as the inverse Gibbs or variational free energy whose maximization results in the minimization of the *Kullback-Leibler* (KL) distance measure between $q_{\mathbf{s}}(\mathbf{s})q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ and $P(\mathbf{s}, \boldsymbol{\theta}|\mathbf{y})$. In order to determine the free distributions $q_{\mathbf{s}}(\mathbf{s})$ and $q_{\boldsymbol{\theta}}(\boldsymbol{\theta})$ that maximize \mathcal{H} , a coordinate ascent algorithm is used that maximizes over one free distribution while keeping the other fixed, in an alternate manner. Based on the functional derivatives of \mathcal{H} with respect to the free distributions, the update equations in the i th iteration are given as follows

$$\begin{aligned} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}) &\propto P(\boldsymbol{\theta}) e^{\sum_{\mathbf{s}} q_{\mathbf{s}}^{(i-1)}(\mathbf{s}) \ln P(\mathbf{y}|\mathbf{s}, \boldsymbol{\theta})}, \\ q_{\mathbf{s}}^{(i)}(\mathbf{s}) &\propto P(\mathbf{s}) e^{\int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}) \ln P(\mathbf{y}|\mathbf{s}, \boldsymbol{\theta}) d\boldsymbol{\theta}}. \end{aligned} \quad (6)$$

The VB based algorithm described is ensured to converge to a fixed point [10]. However, in general global optimality is not guaranteed.

V. RECEIVER DESIGN BASED ON VB FRAMEWORK

In this section, we present an iterative receiver algorithm comprising of a phase estimator and symbol detector based on the VB framework. In order to arrive at the algorithm, the functional form of the free distributions in (6) has to be derived. The received signal model considered is (1). Additionally, the nonlinearity in the received signal model is simplified by locally linearizing the set of phasors $e^{j\theta_{lk}(n)}$ at time instant n using first order Taylor series approximation; i.e., $e^{j\theta_{lk}(n)} \approx e^{j\hat{\theta}_{lk}(n-1)}(1 + j(\theta_{lk}(n) - \hat{\theta}_{lk}(n-1)))$, where $\{\hat{\theta}_{lk}(n-1)\}$ $l = 1, \dots, N_t$, $k =$

$1, \dots, N_r$ are the phase estimates from the previous time instant $n-1$.

Based on (6), the free distribution for the phase noise samples $q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n))$ is given $\forall l \in \{1, \dots, N_t\}$ in the i th iteration as

$$q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n)) \approx P(\boldsymbol{\theta}(n)|\mathbf{y}(n), \underline{\mathbf{s}}(n)), \text{ where,} \quad (7a)$$

$$\sum_{\substack{\mathbf{s}(n) \in \mathcal{C} \\ \mathbf{s}(n) \neq \mathbf{s}_l(n)}} q_{\mathbf{s}_l}^{(i-1)}(\mathbf{s}(n)) = q_{\mathbf{s}_l}^{(i-1)}(s_l(n)), \quad (7b)$$

$$\sum_{\forall s_l(n) \in \mathcal{C}} q_{\mathbf{s}_l}^{(i-1)}(s_l(n))s_l(n) = \underline{s}_l^{(i-1)}(n), \quad (7c)$$

$$\text{Var}_{q_{\mathbf{s}}}\{s_l\} \triangleq \sum_{\mathbf{s}(n) \in \mathcal{C}} q_{\mathbf{s}}^{(i-1)}(\mathbf{s}(n))s_l(n)s_l(n)^* - \underline{s}_l(n)\underline{s}_l(n)^* \approx 0.$$

where $\underline{\mathbf{s}}(n) \triangleq \{\underline{s}_l^{(i-1)}(n)\}$, $\forall l \in \{1, \dots, N_t\}$ is the set of average symbols transmitted by all transmit antennas, as determined in the i th iteration. We refer readers to Appendix A for the proof of this result. Equation (7a) shows that the free distribution of phase that minimizes the KL distance is the posteriori pdf of the phase noise samples given the observation at time instant n and the soft symbols $\{s_l(n)\}$, $\forall l \in 1, \dots, N_t$ obtained based on the free distribution of $\mathbf{s}(n)$. Note that the soft symbols are computed by first marginalizing the joint posteriori pdf of all symbols to find the posteriori of each symbol at time instant n as in (7b). This is then used to determine the average symbol or a soft symbol as in (7b) and (7c). The approximation of the free distribution in (7a) is tight when $\text{Var}_{q_{\mathbf{s}}}\{s_l\} \approx 0$, which implies that the soft symbols are explicitly treated as the true value of $\mathbf{s}(n)$ for the computation of free distribution of $\boldsymbol{\theta}(n)$.

Further, based on (6), the free distribution of the transmitted symbols $q_{\mathbf{s}}^{(i)}(\mathbf{s}(n))$ is given as in (8a). The constant C_1 in (8a) normalizes the *probability mass function* (pmf), and is independent of the symbols transmitted. It is interesting to note the dependencies of the symbol posteriori pmf are as follows

- The posteriori mass of each symbol depends on the distance of the symbol from the received signal after being de-rotated with the estimated phase noise values $\mathbb{E}_{q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n))}\{\theta_{lk}(n)\}$.
- Elements of the covariance matrix of the phase noise estimates in (8c) are weighted by the magnitude of the symbol transmitted by the antennas.

Note that it is possible to replace the exponent I_2 in (8a) with any other appropriate metric (like the one in derived [3]) for the computation of posteriori symbol pmfs.

The phase noise estimates and the covariance matrix are obtained using an *Extended Kalman Filter* (EKF) [4], [15]. This

$$q_{\mathbf{s}}^{(i)}(\mathbf{s}(n)) = C_1 P(\mathbf{s}(n)) \prod_{k=1}^{N_r} \exp\{-I_2\}, \text{ where,}$$

$$I_2 \triangleq \sum_{k=1}^{N_r} \frac{\left| y_k(n) - \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\hat{\theta}_{lk}(n-1)} \right|^2}{2\sigma_w^2} + \sum_{l=1}^{N_t} \frac{|h_{lk} s_l(n)|^2 M_{lk, lk}(n)}{2\sigma_w^2} + \sum_{l=1}^{N_t} \sum_{\substack{m=1 \\ m \neq l}}^{N_t} \frac{h_{lk} h_{mk}^* s_l(n) s_m(n)^* M_{lk, mk}(n) e^{j(\hat{\Delta}_{lm})}}{2\sigma_w^2},$$

$$\text{where, } \hat{\Delta}_{lm} \triangleq \hat{\theta}_{lk}(n-1) - \hat{\theta}_{mk}(n-1). \quad (8a)$$

$$\hat{\theta}_{lk}(n) = \mathbb{E}_{q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n))}\{\theta_{lk}(n)\} \quad (8b)$$

$$M_{lk, mk}(n) = \mathbb{E}_{q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n))}(\theta_{lk}(n) - \hat{\theta}_{lk}(n-1))(\theta_{mk}(n) - \hat{\theta}_{mk}(n-1)), \quad (8c)$$

is a reasonable choice for the linearized observation model since $q_{\theta}(\boldsymbol{\theta}(n))$ in (7a) is Gaussian, given the joint Gaussianity of $\mathbf{y}(n)$ and $\boldsymbol{\theta}(n)$. Note that it is possible to use other nonlinear filters that may not require linearization of the observation model [4], [15]. The linearized state space model for deriving the EKF is given as

$$y_k(n) = \sum_{l=1}^{N_t} h_{lk}(n) e^{j\hat{\theta}_{lk}(n)} (1 + j(\theta_{lk} - \hat{\theta}_{lk})) \underline{s}_l(n) + w_k(n),$$

$$\theta_{lk}(n) = \theta_l^{[t]}(n-1) + \theta_k^{[r]}(n-1) + \Delta_{l,t}(n) + \Delta_{k,r}(n),$$

$$l = 1, \dots, N_t, k = 1, \dots, N_r, n = 1, \dots, N_f.$$

1) *Algorithm Summary*: Based on the free distributions $q_s(\mathbf{s})$ and $q_{\theta}(\boldsymbol{\theta})$ derived, it is now possible to explain the joint estimator-detector algorithm that naturally falls out of the VB framework as follows

- Initialize $\underline{s}_l(n) \forall l \in \{1, \dots, N_t\}$ before the first iteration. Use $\underline{s}_l(n)$ as the true value of transmitted symbols to compute phase noise estimates $\{\hat{\theta}_{lk}(n)\}$ and the covariance matrix at time instant n based on the posteriori pdf $P(\boldsymbol{\theta}(n)|\mathbf{y}(n), \underline{\mathbf{s}}(n))$.
- Based on the phase estimates and its covariance matrix, compute the posteriori pmf of the symbols from (8a) and then soft/average symbols using (7b) and (7c).
- Iterate between the estimator and the detector till the values of $\{\underline{s}_l(n)\}$ and $\{\hat{\theta}_{lk}(n)\}$, $l = 1, \dots, N_t$, $k = 1, \dots, N_r$, $n = 1, \dots, N_f$ converge. After convergence, the detector computes the final hard decisions as

$$\hat{\mathbf{s}}(n) = \underset{\mathbf{s}(n) \in \mathcal{C}}{\operatorname{argmax}} q_s(\mathbf{s}(n)).$$

2) *A Short Note on Complexity*: Computation of the posteriori symbol probabilities according to (8a) involves all possible combination of symbols transmitted from each antenna. This is known to incur high computational complexity, which can be a limiting factor as in any large antenna systems and large constellations [16], [17]. However, it is possible to use a sphere decoder algorithm [16], [17] that helps to reduce the original signal space to a much smaller subset. Hence the posteriori symbol probabilities are only computed for the smaller subset of symbols as decided by the sphere decoder algorithm.

VI. SIMULATION RESULTS

In this section, we present results that demonstrate the performance of the newly proposed receiver in terms of *symbol error probability* (SEP) versus *signal-to-noise ratio* (SNR) per bit. We simulate a MIMO system with $N_t = 2$ transmit antennas, and $N_r = 2$ receive antennas. The channel is assumed to be Rayleigh fading, i.e., $h_{kl}(n) \sim \mathcal{CN}(0, 1)$ and is assumed to be known (estimated). It is assumed that the 3dB bandwidth of the phase noise process is much higher than the doppler spread of the channel. The data transmitted from each antenna is considered to be uncoded and independent of each other. We consider an equal energy constellation - 16-phase shift keying (PSK) and a non-equal energy constellation - 16-quadrature amplitude modulation (QAM) scheme. The phase noise at all the antennas are samples from a Wiener process with innovation variance $\sigma_{l,t}^2 = \sigma_{k,r}^2 = 0.5 \times 10^{-3} \text{ rad}^2$, which corresponds to a strong phase noise scenario. The performance of the proposed detector is compared to that of the optimal *maximum-likelihood* (ML) detector with perfect phase information. Each data frame is considered to be

2000 symbols long with 5 pilot symbols being transmitted at the beginning of each frame. Further a pilot symbol is also transmitted every 15 data symbols resulting in an overall pilot density of around 7.0%. The number of iterations for estimating phase noise and detecting symbols in a time instant n is fixed to 5. Note that SNR/bit used in the simulations depends on the ratio of total useful signal power from all transmit antennas to the AWGN noise at each receive antenna.

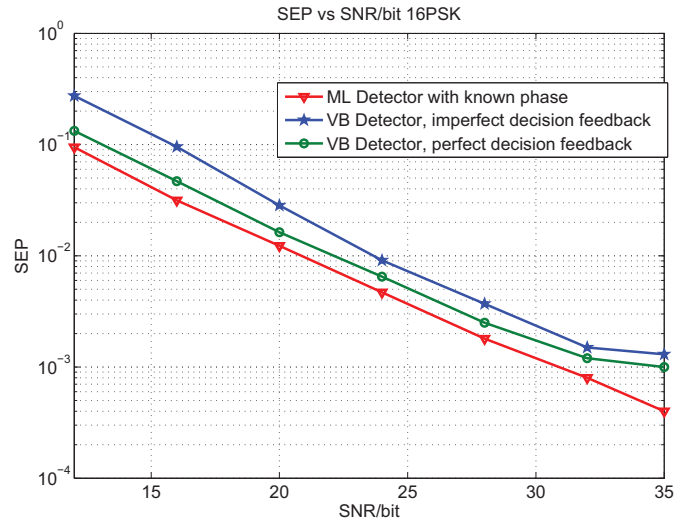


Fig. 1: Comparison of SEP performance between the Optimal ML with perfect phase knowledge and VB based receiver for 16-PSK, $\sigma_{l,t}^2 = \sigma_{k,r}^2 = 0.5 \times 10^{-3} \text{ rad}^2$.

Fig. 1 illustrates the performance of VB detector (with imperfect feedback to estimator) with the optimal ML detector having perfect phase noise information, and perfect decision feedback from the detector to the estimator for 16-PSK modulation scheme. We observe that in the low SNR regime, performance is poor compared to the ML detector with perfect phase information and no improvement is noted even with increase of pilot density to 10%. With increase in SNR, the performance of the detector improves till an error floor is observed at SEP of around 10^{-3} . In Fig. 2, the SEP performance of 16-QAM is presented. It is observed that at low SNR, the performance of the VB detector is similar to that of the ML detector with perfect phase information and the VB detector with perfect decision feedback. As the SNR increases, phase noise becomes more dominant compared to AWGN noise, and the gap between the performance of the two schemes widens till an error floor is observed at SEP of around 4×10^{-3} .

In general, for both constellations, it is unclear whether the error floor observed is a result of the sub-optimality of the detector, or the performance of the estimator. In this regard, it is important to determine a *lower bound* on the performance of the estimator that incorporates soft symbol decisions (where the symbols are assumed to be random variables themselves) and also develop the estimator that achieves that performance, assuming that the bound is tight. Further for both PSK and QAM constellations, we observe that the symbol posteriori distribution becomes multi-peaked in high SNR scenarios, and the symbol decisions from the VB detector depends on the initial value of the estimator-detector algorithm. This is however not observed in low SNR scenarios.

$$I_1 = \frac{-1}{2\sigma_w^2} \sum_{k=1}^{N_r} \sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) \left(y_k(n) y_k(n)^* - \sum_{l=1}^{N_t} \left(|h_{lk} s_l(n)|^2 + \sum_{\substack{m=1 \\ m \neq l}}^{N_t} h_{lk} h_{mk}^* s_l(n) s_m(n)^* e^{j(\theta_{lk}(n) - \theta_{mk}(n))} \right) \right. \\ \left. - y_k(n)^* \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\theta_{lk}(n)} - y_k(n) \sum_{l=1}^{N_t} h_{lk}^* s_l(n)^* e^{-j\theta_{lk}(n)} \right), \quad (\text{A-2})$$

$$= \frac{-1}{2\sigma_w^2} \sum_{k=1}^{N_r} \left(y_k(n) y_k(n)^* + \sum_{l=1}^{N_t} \left(|h_{lk} s_l(n)|^2 + \sum_{\substack{m=1 \\ m \neq l}}^{N_t} h_{lk} h_{mk}^* s_l(n) s_m(n)^* e^{j(\theta_{lk}(n) - \theta_{mk}(n))} \right) - y_k(n)^* \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\theta_{lk}(n)} \right. \\ \left. - y_k(n) \sum_{l=1}^{N_t} h_{lk}^* s_l(n)^* e^{-j\theta_{lk}(n)} - \sum_{l=1}^{N_t} \text{Var}\{s_l\} \right), \quad (\text{A-3})$$

$$\text{Put } \text{Var}\{s_l\} = \sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) s_l(n) s_l(n)^* - s_l(n) s_l(n)^* \approx 0.$$

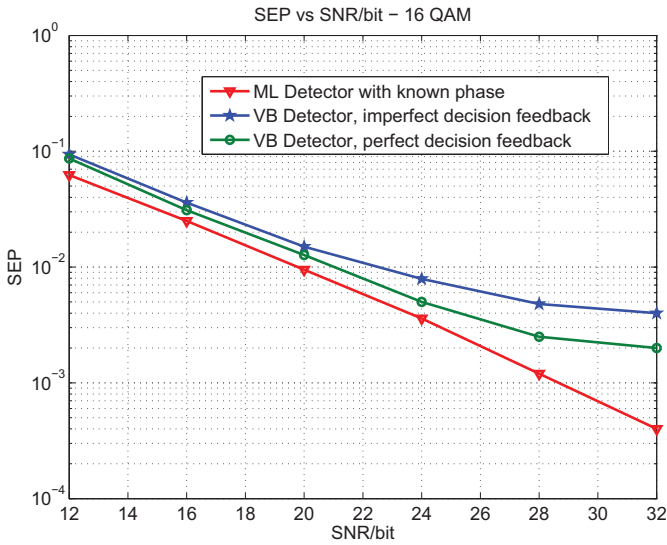


Fig. 2: Comparison of SEP performance between the Optimal ML with perfect phase knowledge and VB based receiver for 16-QAM, $\sigma_{i,t}^2 = \sigma_{k,r}^2 = 0.5 \times 10^{-3}$ rad².

VII. CONCLUSIONS

In this paper, we motivate the problem of receiver design in the presence of time varying random oscillator phase noise in a MIMO system. We observe that there are interesting scenarios where the time varying oscillator phase noise varies much faster than a time varying channel. In such scenarios, the impact of phase noise need not be moved into the channel matrix for the purpose of complexity reduction. Also we observe that the phase noise in MIMO impacts both phase and amplitude of the signal, unlike in SISO systems. Then using the VB framework, an iterative estimator-detector algorithm is developed for data detection in the presence of Wiener phase noise. The SEP of the newly proposed iterative scheme is compared with that of the optimal ML detector with perfect phase information and the VB detector with perfect decision feedback for 16-PSK and 16-QAM modulation schemes. In both modulation schemes, the gap of performance of the proposed technique with respect to the optimal ML detector with perfect phase information increases with SNR till an error floor (due to phase noise) is observed.

APPENDIX A

DERIVATION OF THE POSTERIORI PHASE PROBABILITIES

From (6), the free distribution of $\theta(n)$ is as follows

$$q_\theta^{(i)}(\theta(n)) = CP(\theta(n)) \times \exp \left\{ \underbrace{\sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) \ln P(\mathbf{y}(n)|\mathbf{s}(n), \theta(n))}_{I_1} \right\}.$$

Simplifying the exponent I_1 , we have

$$I_1 = \sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) \ln \prod_{k=1}^{N_r} P(y_k(n)|\mathbf{s}(n), \theta(n)), \\ = \sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) \sum_{k=1}^{N_r} \ln P(y_k(n)|\mathbf{s}(n), \theta(n)), \\ = \frac{-1}{2\sigma_w^2} \sum_{k=1}^{N_r} \sum_{\mathbf{s}(n) \in \mathcal{C}} q_s^{(i-1)}(\mathbf{s}(n)) \\ \times \left| y_k(n) - \sum_{l=1}^{N_t} s_l(n) e^{j\theta_{lk}(n)} h_{lk} \right|^2. \quad (\text{A-1})$$

Using the simplification for I_1 from (A-2), the posteriori phase pdf is as follows

$$q_\theta^{(i)}(\theta(n)) = CP(\theta(n)) \\ \times \exp \left\{ \sum_{k=1}^{N_r} \frac{\left| y_k(n) - \sum_{l=1}^{N_t} s_l(n) e^{j\theta_{lk}(n)} \right|^2}{2\sigma_w^2} \right\} \\ = CP(\theta(n)) P(\mathbf{y}(n)|\theta(n), \underline{\mathbf{s}}(n)), \quad (\text{A-4}) \\ = CP(\mathbf{y}(n), \theta(n)|\underline{\mathbf{s}}(n)). \quad (\text{A-5})$$

The constant C is pdf normalizing factor such that $q_\theta^{(i)}(\theta(n))$ integrates to unity. Its value can be determined as

$$C = P(\mathbf{y}(n)|\underline{\mathbf{s}}(n)).$$

$$\begin{aligned}
I_2 &= \frac{-1}{2\sigma_w^2} \sum_{k=1}^{N_r} \mathbb{E}_{q_{\theta}^{(i)}(\boldsymbol{\theta}(n))} \left(y_k(n) - \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\theta_{lk}(n)} \right) \left(y_k(n)^* - \left(\sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\theta_{lk}(n)} \right)^* \right), \\
&= \sum_{k=1}^{N_r} \frac{-1}{2\sigma_w^2} \left\{ y_k(n) y_k(n)^* + \sum_{l=1}^{N_t} \left(|h_{lk} s_l(n)|^2 \left(1 + \mathbb{E}_{q_{\theta}^{(i)}(\boldsymbol{\theta}(n))} (\theta_{lk}(n) - \hat{\theta}_{lk}(n-1))^2 \right) \right. \right. \\
&\quad \left. \left. + \sum_{\substack{m=1 \\ m \neq l}}^{N_t} h_{lk} h_{mk}^* s_l(n) s_m(n)^* \left(e^{j(\hat{\theta}_{lk}(n-1) - \hat{\theta}_{mk}(n-1))} \left(1 + \mathbb{E}_{q_{\theta}^{(i)}(\boldsymbol{\theta}(n))} (\theta_{lk}(n) - \hat{\theta}_{lk}(n-1)) (\theta_{mk}(n) - \hat{\theta}_{mk}(n-1)) \right) \right) \right) \right. \\
&\quad \left. - y_k(n)^* \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\hat{\theta}_{lk}(n-1)} - y_k(n) \sum_{l=1}^{N_t} h_{lk}^* s_l(n)^* e^{-j\hat{\theta}_{lk}(n-1)} \right\}, \tag{B-2} \\
&= \sum_{k=1}^{N_r} - \frac{\left| y_k(n) - \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\hat{\theta}_{lk}(n-1)} \right|^2}{2\sigma_w^2} - \sum_{l=1}^{N_t} \frac{|h_{lk} s_l(n)|^2 M_{lk, lk}(n)}{2\sigma_w^2} - \sum_{l=1}^{N_t} \sum_{\substack{m=1 \\ m \neq l}}^{N_t} \frac{h_{lk} h_{mk}^* s_l(n) s_m(n)^* M_{lk, mk}(n) e^{j(\hat{\Delta}_{lm})}}{2\sigma_w^2}, \\
&\quad \text{where, } \hat{\Delta}_{lm} \triangleq \hat{\theta}_{lk}(n-1) - \hat{\theta}_{mk}(n-1). \tag{B-3}
\end{aligned}$$

Plugging the value of C in (A-4), the free distribution of $\boldsymbol{\theta}(n)$ can be obtained as

$$q_{\theta}^{(i)}(\boldsymbol{\theta}(n)) = P(\boldsymbol{\theta}(n) | \mathbf{y}(n), \underline{\mathbf{s}}(n)), \text{ where,} \tag{A-6}$$

$$\sum_{\substack{s_k \in \mathcal{S} \\ \mathbf{s}(n) \in \mathcal{C}}} q_{\mathbf{s}}^{(i-1)}(\mathbf{s}(n)) s_l(n) = \underline{s}_l^{(i-1)}(n). \tag{A-7}$$

APPENDIX B

DERIVATION OF THE POSTERIORI SYMBOL PROBABILITIES

From (6), the free distribution of $\mathbf{s}(n)$ is as follows

$$\begin{aligned}
q_{\mathbf{s}}^{(i)}(\mathbf{s}(n)) &= C_1 P(\mathbf{s}(n)) \\
&\times \exp \left\{ \underbrace{\int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n)) \ln P(\mathbf{y}(n) | \mathbf{s}(n), \boldsymbol{\theta}(n)) d\boldsymbol{\theta}(n)}_{I_2} \right\},
\end{aligned}$$

The exponent I_2 can be simplified as follows

$$\begin{aligned}
I_2 &= \int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n)) \ln \prod_{k=1}^{N_r} P(y_k(n) | \mathbf{s}(n), \boldsymbol{\theta}(n)) d\boldsymbol{\theta}(n), \\
&= \int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n)) \sum_{k=1}^{N_r} \ln P(y_k(n) | \mathbf{s}(n), \boldsymbol{\theta}(n)) d\boldsymbol{\theta}(n), \\
&= \frac{-1}{2\sigma_w^2} \sum_{k=1}^{N_r} \int_{\boldsymbol{\theta}} q_{\boldsymbol{\theta}}^{(i)}(\boldsymbol{\theta}(n)) \\
&\quad \times \left| y_k(n) - \sum_{l=1}^{N_t} h_{lk} s_l(n) e^{j\theta_{lk}(n)} \right|^2 d\boldsymbol{\theta}(n), \tag{B-1}
\end{aligned}$$

Hence we obtain the result in (8a).

REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to SpaceTime Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [2] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers, Synchronization, Channel Estimation, and Signal Processing*. Wiley, 1998.
- [3] R. Krishnan, H. Mehrpouyan, T. Eriksson, T. Svensson, "Optimal and Approximate Methods for Detection of Uncoded Data with Carrier Phase Noise," Accepted for IEEE Globecom 2011, Houston.
- [4] M. Nissila, S. Pasupathy, "Adaptive iterative detectors for phase-uncertain channels via variational bounding," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 716-725, Mar. 2009.
- [5] G. Colavolpe, A. Barbieri, and G. Caire, "Algorithms for iterative decoding in the presence of strong phase noise," *IEEE Journal Selected Areas Commun.*, vol. 23, no. 9, pp. 1748-1757, Sept. 2005
- [6] Y. Zhang; L. Huaping, "MIMO-OFDM Systems in the Presence of Phase Noise and Doubly Selective Fading," *IEEE Trans. Veh. Tech.*, vol.56, no.4, pp. 2277-2285, Jul. 2007.
- [7] D. S. Baum and H. Bölcskei, "Information-theoretic analysis of MIMO channel sounding," *IEEE Trans. Info. Theory*, 2011, to appear.
- [8] P. Almers, S. Wyne, F. Tufvesson, A. F. Molisch, "Effect of random walk phase noise on MIMO measurements," *IEEE 61st VTC 2005-Spring*, pp. 141- 145, 30 May-Jun. 2005.
- [9] K. Huber, S. Haykin, "Improved bayesian MIMO channel tracking for wireless communications: incorporating a dynamical model," *IEEE Trans. Wireless Comm.*, vol.5, no.9, pp.2458-2466, Sep. 2006.
- [10] M. J. Beal, "Variational algorithms for approximate Bayesian inference," Ph.D. dissertation, Univ. Coll. London, 1998.
- [11] M. J. Wainwright, M. I. Jordan, *Graphical Models, Exponential Families, and Variational Inference*. Hanover, MA, USA: Now Publishers Inc., 2008.
- [12] L. D. Darryl, T. J. Lim, "The Variational Inference Approach to Joint Data Detection and Phase Noise Estimation in OFDM," *IEEE Trans. Sig. Process.*, vol. 55, no. 5, pp. 1862-1874, May 2007.
- [13] T. Schenk, *RF Imperfections in High-rate Wireless Systems: Impact and Digital Compensation*, Springer Publishing Company, 2008.
- [14] A. Demir, A. Mehrotra, J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Trans. Circ. Sys. I: Fundamental Theory and Applications*, vol.47, no.5, pp.655-674, May 2000.
- [15] S. M. Kay, *Fundamentals of Statistical Signal Processing, Estimation Theory*, vol. 1, Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [16] E. Agrell, T. Eriksson, T. A. Vardy, K. Zeger, "Closest point search in lattices," *IEEE Trans. Info. Theory*, vol.48, no.8, pp. 2201- 2214, Aug 2002.
- [17] C. Xiong; H. Wang; X. Zhang; J. Wei; Chao-jing Tang, "Low Complexity Variational Bayes Iterative Receiver for MIMO-OFDM Systems," *IEEE Int. Conf. Comm., 2009. ICC '09.*, vol., no., pp.1-5, 14-18 June 2009.