Joint Estimation of Channel and Oscillator Phase Noise in MIMO Systems

Hani Mehrpouyan, *Member, IEEE*, Ali A. Nasir, *Student Member, IEEE*, Steven D. Blostein, *Senior Member, IEEE*, Thomas Eriksson, George K. Karagiannidis, *Senior Member, IEEE*, and Tommy Svensson, *Senior Member, IEEE*

Abstract—Oscillator phase noise limits the performance of high speed communication systems since it results in time varying channels and rotation of the signal constellation from symbol to symbol. In this paper, joint estimation of channel gains and Wiener phase noise in multi-input multi-output (MIMO) systems is analyzed. The signal model for the estimation problem is outlined in detail and new expressions for the Cramér-Rao lower bounds (CRLBs) for the multi-parameter estimation problem are derived. A data-aided least-squares (LS) estimator for jointly obtaining the channel gains and phase noise parameters is derived. Next, a decision-directed weighted least-squares (WLS) estimator is proposed, where pilots and estimated data symbols are employed to track the time-varying phase noise parameters over a frame. In order to reduce the overhead and delay associated with the estimation process, a new decision-directed extended Kalman filter (EKF) is proposed for tracking the MIMO phase noise throughout a frame. Numerical results show that the proposed LS, WLS, and EKF estimators' performances are close to the CRLB. Finally, simulation results demonstrate that by employing the proposed channel and time-varying phase noise estimators the bit-error rate performance of a MIMO system can be significantly improved.

Index Terms—Channel estimation, Cramér-Rao lower bound (CRLB), extended Kalman filter (EKF), multi-input multi-output (MIMO), weighted least squares (WLS), Wiener phase noise.

I. INTRODUCTION

A. Motivation and Literature Survey

WIRELESS communication links are expected to carry ever higher rates over the available bandwidth. The extensive research in the field of multi-input multi-output (MIMO)

systems has shown that such systems are capable of significantly enhancing bandwidth efficiency of wireless systems [1], [2]. However, achieving synchronous transmission in today's high speed wireless systems is a challenging task, since many rapidly varying synchronization parameters need to be simultaneously and jointly estimated at the receiver [3]. Therefore, accurate and efficient algorithms that enable synchronous high-speed communication are of broad interest [4].

Phase noise, which is present in wireless communication systems due to imperfect oscillators [5]–[12], is greatly detrimental to synchronization unless its parameters are accurately estimated and compensated [3]. The effect of phase noise on the performance of wireless communication systems is more pronounced at higher carrier frequencies [4]. In addition, motivated by the large available bandwidth in the E-band (60–80 GHz) extensive research has been recently carried out on efficient algorithms that are capable of outperforming traditional phase noise tracking schemes, e.g., those based on the phase-locked loop (PLL). In the case of MIMO systems, each transmit and receive antenna may be equipped with an independent oscillator. For example, in the case of line-of-sight (LoS) MIMO systems[1], a single oscillator cannot be used for all the transmit or receive antennas since the antennas need to be placed far apart from one another [13][2]. Similarly, in multiuser MIMO or space division multiple access (SDMA) systems, multiple users with independent oscillators transmit their signals to a common receiver [15]. Thus, in order to comprehensively address the problem of phase noise mitigation in MIMO systems, algorithms that can jointly estimate multiple phase noise parameters at the receiver are of particular interest [13], [15].

In single carrier communication systems, phase noise is multiplicative and results in a rotation of the signal constellation from symbol to symbol and erroneous data detection [3]. The Cramér-Rao lower bounds (CRLB) and algorithms for estimation of phase noise in single-input single-output (SISO) systems are extensively and thoroughly analyzed in [3], [16]–[30]. However, these results are not applicable to MIMO systems, where a received signal may be affected by multiple phase noise parameters that need to be jointly estimated at the receiver [13], [31]–[33]. As a matter of fact, for SISO systems, Kalman filter-based methods have been effectively applied in [16], [20], [25], [29] for signal detection in the presence of phase noise. However, as stated previously these approaches are not applicable to the case of MIMO systems and the results in [16], [20], [25],

---

[1]LoS MIMO has been effectively demonstrated for microwave backhauling by Ericsson AB.

[2]For a 4×4 LoS MIMO system operating at 10 GHz and with a transmitter and receiver distance of 2 km, the optimal antenna spacing is 3.8 m at both the transmitter and receiver [14].

and [29] are focused on signal detection and do not analyze or investigate the estimation performance of Kalman filter based phase noise mitigation methods.

The effect of phase noise on the capacity and performance of space division multiplexing (SDM) MIMO communication systems has been analyzed in [31]–[33], where it is demonstrated that phase noise can greatly limit the performance of multi-antenna systems. In [31], the effect of oscillator configuration at the transceiver antennas and the resulting phase noise on a MIMO beamforming system is analyzed in detail. In [32], [33], it is demonstrated that imperfect knowledge of phase noise and channel have a considerable impact on the capacity of MIMO systems. However, the impact of imperfect knowledge of phase noise and channel on the performance of a MIMO system for different numbers of antennas, modulation schemes, and phase noise variances is not investigated. In [13], pilot-aided estimation of phase noise in a MIMO system is investigated, where a Wiener filtering approach is applied to estimate the phase noise parameters corresponding to the transmit and receive antennas. However, the scheme in [13] is based on the assumption that the MIMO channel is *perfectly* known at the receiver and it requires that while one antenna transmits its pilots, the remaining antennas stay silent, which is bandwidth inefficient and results in significant overhead. In addition, in [13], the CRLB for the estimation problem is not derived and the performance of the proposed MIMO phase noise estimator is not investigated. To the best of the authors' knowledge, a complete analysis of the joint estimation of channel and phase noise parameters in MIMO systems has not been addressed in the literature to date.

In orthogonal frequency division multiplexing (OFDM) phase noise is convolved with the data symbols [7], [34]–[36]. Therefore, the effect of phase noise can be partitioned into a multiplicative part and an additive part that results in inter-carrier interference (ICI) and significant performance loss [7]. Since in OFDM systems, the multiplicative part of phase noise affects all subcarriers similarly [35], it is referred to as the common phase error (CPE) and can be easily compensated by a phase rotation as shown in [7], [34], [36], [37]. On the other hand, the additive part of phase noise is more challenging to mitigate and considerable research has been carried out to analyze and reduce its resulting ICI in SISO-and MIMO-OFDM systems [7], [35], [38]. More specifically, the algorithms in [7], [36], [39]–[41] are only applicable to SISO-OFDM systems and do not provide any means of estimating or tracking multiple phase noise parameters. In [42], it is shown that channel and phase noise need to be *jointly* estimated at a SISO-OFDM receiver and new algorithms for obtaining them are presented. In order to further improve system performance, an approach for joint estimation and suppression of CPE and ICI in SISO-OFDM systems based on the variational inference approach have been presented in [43], where it is discussed that in most practical scenarios of interest, the phase noise process varies much more quickly than the channel and, as a result, the effect of phase noise cannot be mitigated using a simple training approach. However, the results in [42], [43] are only applicable to SISO-OFDM systems, the derived CRLB for channel estimation [42] does not take the effect of phase noise into account, and the estimators in [42] can be only applied to wireless systems where the receiver is equipped with a PLL. Furthermore, the iterative algorithms in [43] can

potentially result in unwanted delay and overhead in high speed communication systems.

Even though the results in [35] provide schemes for mitigating the effect of phase noise induced ICI in MIMO-OFDM systems, they present no means of estimating the CPE or the multiplicative phase noise affecting these communication systems. Algorithms for CPE estimation in MIMO-OFDM systems have been proposed in [34], [37], [38]. However, these schemes are based on the assumption that the MIMO channels are known and are limited to scenarios where a single oscillator is used at all the transmit or receive antennas. As a result, the approaches in [34], [37], [38] cannot track multiple channels and phase noise parameters and cannot be applied to various MIMO systems. Moreover, in [34], [37], [38], no specific performance bound, e.g., CRLB, for the estimation of channel and phase noise parameters is derived. In [15], a new algorithm for estimation of the CPE in SDMA MIMO-OFDM systems has been proposed. Even though the maximum a posteriori estimator in [15] can track multiple phase noise parameters, it has a very high computational complexity, it is based on the assumption that the MIMO channels are perfectly known, and the performance of the proposed estimators is only verified for low-to-moderate phase noise variances. Finally, in [15], the CRLB for estimation of multiple phase noise parameters in SDMA MIMO systems is not derived.

It is important to note that compared to single carrier systems, phase noise deteriorates the performance of OFDM systems more significantly [44]. This sensitivity to phase noise in OFDM systems is even more severe as the constellation size and number of subcarriers increases [44]. Therefore, application of single carrier systems to very high speed communication links may be advantageous. Moreover, application of single carrier SDM instead of OFDM in wireless communication systems that operate in frequency non-selective fading channels may result in reduced overhead and cost, since OFDM signalling requires additional signal processing at the transmitter and receiver (a fast Fourier transform (FFT) and an inverse FFT at the receiver and transmitter, respectively) [45], necessitates more overhead due to the cyclic prefix [45], and requires linear amplifiers [46]. For example, in the case of high speed LoS microwave backhaul links (backhaul networks connect cellular base stations to the core network), the wireless channel is not frequency selective [47]–[49] and, therefore, single carrier SDM is used instead of OFDM [47]. Moreover, OFDM may not be suitable for wireless systems that operate in nonlinear channels due to OFDM signals' amplitude variations and high peak-to-average power ratio (PAPR) [46]. An example can be found in satellite communication links, where use of high power amplifiers results in significant non-linearity in the wireless channel [46, p. 383]. Finally, single carrier systems are considered to be advantageous in the E-band due to their low PAPR (at very high carrier frequencies the dynamic range of power amplifiers is limited) and their better performance when using high rate or weak error correcting codes [50, p. 261].

B. Contributions

In this paper, joint estimation of multiple phase noise parameters and channel gains in a single carrier SDM MIMO system equipped with N_t transmit and N_r receive antennas is analyzed. The system model for the estimation problem is formulated

in detail and new CRLBs for the multiple parameter estimation problem corresponding to *online* processing of the received signal are derived. A data-aided least squares (LS) estimator for jointly obtaining the channel gains and phase noise values is derived. Next, the pilot and estimated data symbols in combination with a *decision-directed*[3] weighted least-squares (WLS) estimator are used to track the time-varying phase noise processes over a frame. In order to reduce overhead and delay associated with the estimation process, a decision-directed extended Kalman filter (EKF) is also proposed. The performance of the proposed LS, WLS, and EKF based channel and phase noise estimators is shown to be close to the derived CRLBs over a wide range of signal-to-noise ratio (SNR) values. Moreover, simulation results demonstrate that by employing the proposed channel and phase noise estimators the bit-error rate (BER) performance of a MIMO system can be significantly improved in the presence of time-varying phase noise. Finally, the effect of number of antennas, modulation scheme, and phase noise variance on the performance of MIMO systems with imperfect knowledge of channel and phase noise is investigated[4]. The contributions of this paper can be summarized as follows:

- The joint estimation of channel gains and phase noise in an $N_r \times N_t$ MIMO system is parameterized and new CRLBs for the multi-parameter estimation problem in cases of both data-aided and decision-directed estimation are derived. The CRLBs are then used as a benchmark for the performance of the proposed estimators and are also applied to quantitatively determine the effect of unknown phase noise on channel estimation accuracy and vice versa.

- Algorithms for estimating and tracking the unknown channel gains and time-varying phase noise, respectively, throughout a frame are proposed. A data-aided LS estimator for jointly obtaining the MIMO channel and phase noise parameters is proposed. Next, novel WLS and EKF based estimators are proposed that are shown to accurately track the phase noise over a frame and reach the derived CRLB. A complexity analysis is carried out to show that the proposed EKF can efficiently track multiple noisy and time-varying phase shifts in a MIMO system.

- Extensive simulations are carried out that investigate the performance of MIMO systems in the presence of imperfectly estimated channels and phase noises for different phase noise variances, modulations, synchronization overheads, channel conditions, and Doppler rates. These simulations demonstrate that application of the proposed channel and phase noise estimators can significantly improve the performance of MIMO systems.

C. Organization

The remainder of the paper is organized as follows: in Section II the phase noise model and MIMO framework used throughout the paper are outlined, Section III derives the new CRLBs for both cases of data-aided and decision-directed estimation, Section IV presents the novel channel and phase noise estimation algorithms while Section V provides numerical and simulation results that examine the performance of MIMO systems in the presence of estimated channel and phase noise.

---

[3]For decision-directed, the prior pilot and estimated data symbols are used to estimate the current symbol's phase noise parameters.

[4]This paper is partly presented at SPAWC 2012 [51].

Finally, Section VI concludes the paper and summarizes its key findings.

Notation

Superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote conjugate, conjugate transpose, and transpose operators, respectively. Bold face small letters, e.g., $\mathbf{x}$, are used for vectors, bold face capital alphabets, e.g., $\mathbf{X}$, are used for matrices, and $[\mathbf{X}]_{x,y}$ represents the entry in row x and column y of $\mathbf{X}$. $\mathbf{I}_{X \times X}$, $\mathbf{0}_{X \times X}$, and $\mathbf{1}_{X \times X}$, denote the $X \times X$ identity, all zero, and all 1 matrices, respectively. $\odot$ stands for Schur (element-wise) product, $|\cdot|$ is the absolute value operator, $\angle x$ returns the phase of complex variable x, $|\mathbf{x}|$ denotes the element-wise absolute value of a vector $\mathbf{x}$, $\mathrm{diag}(\mathbf{x})$ is used to denote a diagonal matrix, where the diagonal elements are given by vector $\mathbf{x}$. $\mathrm{diag}(\mathbf{X})$ is used to denote the diagonal elements of matrix $\mathbf{X}$. $\mathbb{E}[\cdot]$ denotes the expected value of the argument, and $\Re\{\cdot\}$ and $\Im\{\cdot\}$ are the real and imaginary parts of a complex quantity, respectively. Finally, $\mathcal{N}\left(\mu, \sigma^2\right)$ and $\mathcal{CN}\left(\mu, \sigma^2\right)$ denote real and complex Gaussian distributions with mean μ and variance σ^2, respectively.

II. SYSTEM MODEL

A point-to-point MIMO system with N_t and N_r transmit and receive antennas, respectively, is considered (see Fig. 1). As shown in Fig. 2 each frame of length L_f symbols is assumed to consist of a training sequence (TS) of length L_t symbols, *data* symbols, and *pilot* symbols that are transmitted every L_p symbol interval. In this paper, the following set of assumptions is adopted:

A1. The pilot symbols are assumed to be *known* at the receiver. Moreover, it is assumed that all transmit antennas simultaneously broadcast mutually orthogonal TSs of length $L_t = N_t$ to the receiver.

A2. In order to ensure generality, each transmit and receive antenna is assumed to be equipped with an independent oscillator as depicted in Fig. 1. This ensures that the system model is in line with previous work in [13] and is also applicable to various MIMO scenarios, e.g., LoS MIMO and SDMA MIMO systems.

A3. The analyses in this paper are based on the assumption of Quasi-static and frequency-flat fading channels, where the channel gains are assumed to remain constant over a frame, i.e., the channel gains are modeled as *unknown deterministic* parameters over a frame. Nevertheless, in Section V, the performances of the proposed channel and phase noise tracking schemes in the presence of time-varying channels are investigated.

A4. The time-varying phase noise process is modeled as a random-walk or Wiener model. It should be noted that phase noise is assumed to evolve much more slowly than the symbol rate. Therefore, phase noise is assumed to not change during the duration of a symbol but to change from symbol to symbol.

A5. Perfect timing and frame synchronization is assumed, which can be achieved by standard frame synchronization algorithms using a timing feedback loop [18].

Note that Assumptions A3, A4, and A5 are in line with previous phase noise estimation algorithms in SISO and MIMO systems in [3], [13], [17], [18], [52]–[54]. Moreover, Assumption A3 is reasonable in many practical scenarios, e.g., in LoS
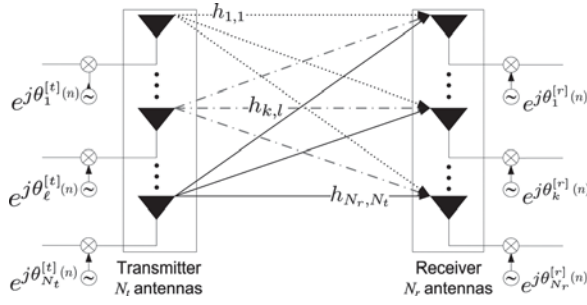
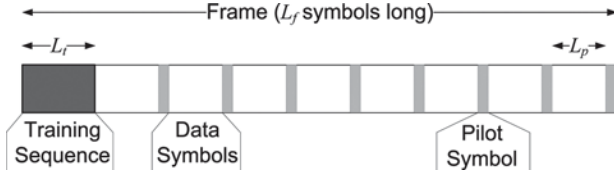Fig. 1. System model for a point-to-point MIMO systems.

Fig. 2. Timing diagram for transmission of training, pilot, and data symbols within a frame.

MIMO systems applied to microwave backhaul [47] and satellite communication links [46], where the channel gains vary much more slowly than the phase noise process. More importantly, unlike the results in [3], [13], [17], [18], [52]–[54], which assume that the channel gains are estimated and equalized before phase noise estimation, in this paper we jointly estimate the MIMO channel gains and phase noise parameters. Note that even though the analyses in this paper are based on the assumption of quasi-static fading channels, in Section V, it is demonstrated that by selecting an appropriate synchronization overhead, the proposed estimators can accurately track MIMO channels and phase noise processes in the presence of time-varying channels with different Doppler rates.

The discrete-time baseband received signal model at the kth antenna of the MIMO receiver is given by[5]

$$
\begin{aligned}
y_k(n) &= \sum_{\ell=1}^{N_t} e^{j\theta_k^{[r]}(n)} h_{k,\ell} e^{j\theta_\ell^{[t]}(n)} s_\ell(n) + w_k(n) \\
&= \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} s_\ell(n) + w_k(n), \quad k=1,\ldots,N_r
\end{aligned}
$$
(1)

where

- $s_\ell(n)$ is the nth, for $n = 1, \ldots, L_f$, M-ary modulated transmitted symbol that corresponds to the ℓth transmit antenna and consists of both pilots and data symbols;
- $h_{k,\ell}$ is the quasi-static *unknown* channel gain from the ℓth transmit to the kth receive antenna, which is assumed to be constant over the length of a frame and to be distributed as a complex Gaussian random variable, i.e., $h_{k,\ell} \sim \mathcal{CN}(\mu_{h_{k,\ell}}, \sigma_{h_{k,\ell}}^2)$ from frame to frame;
- $\alpha_{k,\ell} \triangleq |h_{k,\ell}|$ is the channel gain from the ℓth transmit to the kth receive antenna;

[5]Throughout this paper indices $\ell = 1, \ldots, N_t$, $k = 1, \ldots, N_r$, and $n = 1, \ldots, L_f$ are used to denote transmit antennas, receive antennas, and symbols, respectively.

- $\theta_\ell^{[t]}(n)$ and $\theta_k^{[r]}(n)$ correspond to the nth sample of the phase noise at the ℓth transmit and kth receive antenna, respectively;
- $\beta_{k,\ell}(n) \triangleq \theta_\ell^{[t]}(n) + \theta_k^{[r]}(n) + \angle h_{k,\ell}$ denotes the overall phase shift from the oscillator and channel corresponding to the ℓth transmit and kth receive antenna; and
- $w_k(n)$ is the zero-mean complex additive white Gaussian noise (AWGN) at the kth receive antenna, i.e., $w_k(n) \sim \mathcal{CN}(0, \sigma_w^2)$. Note that the AWGN variance, σ_w^2, is assumed to be known since it can be estimated at the receiver [55].

The discrete time phase noise model in (1) is motivated by the results in [3], [7], [13]. More importantly, for free-running oscillators, it is found that the phase noise process can be modeled as a Brownian motion or Wiener process [5]–[10]. Therefore, $\theta_\ell^{[t]}$ and $\theta_k^{[r]}$, for $\ell = 1, \ldots, N_t$ and $k = 1, \ldots, N_r$, are given by [5]–[10]

$$
\begin{aligned}
\theta_\ell^{[t]}(n) &= \theta_\ell^{[t]}(n-1) + \Delta_\ell^{[t]}(n) \\
\theta_k^{[r]}(n) &= \theta_k^{[r]}(n-1) + \Delta_k^{[r]}(n)
\end{aligned}
$$
(2)

where the phase innovations for the ℓth transmit and kth receive antennas, $\Delta_\ell^{[t]}(n)$ and $\Delta_k^{[r]}(n)$, respectively, are assumed to be white real Gaussian processes with $\Delta_\ell^{[t]}(n) \sim \mathcal{N}\left(0, \sigma_{\Delta_\ell^{[t]}}^2\right)$ and $\Delta_k^{[r]}(n) \sim \mathcal{N}\left(0, \sigma_{\Delta_k^{[r]}}^2\right)$. The variances of the innovations or the *phase noise rate* at the ℓth and kth transmit and receive antennas, $\sigma_{\Delta_\ell^{[t]}}^2$ and $\sigma_{\Delta_k^{[r]}}^2$, respectively, are given by [5]–[10], [38]

$$
\begin{aligned}
\sigma_{\Delta_\ell^{[t]}}^2 &= 2\pi c_\ell^{[t]} T_s, \\
\sigma_{\Delta_k^{[r]}}^2 &= 2\pi c_k^{[r]} T_s
\end{aligned}
$$
(3)

where the constants $c_\ell^{[t]}$ and $c_k^{[r]}$ denote the one-sided 3-dB bandwidth of the Lorentzian spectrum of the oscillators at the ℓth and kth transmit and receive antennas, respectively, and T_s is the sampling time. As shown in [11], [12], in practice, the phase noise innovation variance is small, e.g., using the measurement results in [12, Fig. 16], and [12, Eq. (10)] for a free-running oscillator operating at 2.8 GHz with $T_s = 10^{-6}$ sec, the phase noise rate is calculated to be $\sigma_\Delta^2 = 10^{-4}$ rad^2. Finally, throughout this paper it is assumed that $\sigma_{\Delta_\ell^{[t]}}^2$ and $\sigma_{\Delta_k^{[r]}}^2$ are known at the receiver, given that they are dependent on the oscillator properties.

Equation (1) can be written in vector and matrix form as

$$
\begin{aligned}
\mathbf{y}(n) &= \mathbf{\Theta}^{[r]}(n)\mathbf{H}\mathbf{\Theta}^{[t]}(n)\mathbf{s}(n) + \mathbf{w}(n) \\
&= \underbrace{\left(\mathbf{A} \odot e^{j\mathbf{B}(n)}\right)}_{\triangleq \mathbf{P}(n)} \mathbf{s}(n) + \mathbf{w}(n)
\end{aligned}
$$
(4)

where

- $\mathbf{y}(n) \triangleq [y_1(n), \ldots, y_{N_r}(n)]^T$;
- $\mathbf{s}(n) \triangleq [s_1(n), \ldots, s_{N_t}(n)]^T$;
- $\mathbf{H} \triangleq [\mathbf{h}_1, \ldots, \mathbf{h}_{N_r}]^T$;
- $\mathbf{h}_k \triangleq [h_{k,1}, \ldots, h_{k,N_t}]^T$, for $k = 1, \ldots, N_r$;
- $\mathbf{\Theta}^{[r]}(n) \triangleq \operatorname{diag}\left(e^{j\theta_1^{[r]}(n)}, \ldots, e^{j\theta_{N_r}^{[r]}(n)}\right)$ is an $N_r \times N_r$ matrix;

- $\mathbf{\Theta}^{[t]}(n) \triangleq \mathrm{diag}\left(e^{j\theta_1^{[t]}(n)}, \ldots, e^{j\theta_{N_t}^{[t]}(n)}\right)$ is an $N_t \times N_t$ matrix;
- $\mathbf{B}(n) \triangleq \left[\boldsymbol{\beta}_1(n), \ldots, \boldsymbol{\beta}_{N_r}(n)\right]^T$ is an $N_r \times N_t$ matrix;
- $\boldsymbol{\beta}_k(n) \triangleq \left[\beta_{k,1}(n), \ldots, \beta_{k,N_t}(n)\right]^T$;
- $\mathbf{A} \triangleq \left[\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_{N_r}\right]^T$ is the $N_r \times N_t$ channel gain matrix;
- $\boldsymbol{\alpha}_k \triangleq \left[\alpha_{k,1}, \ldots, \alpha_{k,N_t}\right]^T$; and
- $\mathbf{w}(n) \triangleq \left[w_1(n), \ldots, w_{N_r}(n)\right]^T$.

The following remarks are in order.

Remark 1: Given that both the channel gains, $\mathbf{A}$, and phase fluctuations, $\mathbf{B}(n)$, in (4) are assumed to be unknown, the TS at the beginning of each frame is used to jointly estimate a total of $2N_t N_r$ parameters at the receiver. Next, since in most practical scenarios of interest, the channel gains vary much more slowly compared to the phase noise processes [43], the pilot and estimated data symbols are used to only estimate the $N_t N_r$ phase noise parameters, $\mathbf{B}(n)$, over the frame.

Remark 2: It is a well-known fact that the Bayesian CRLB (BCRLB) [56, p. 84] is better suited for determining the lower bound on the estimation accuracy of random parameters. However, deriving the exact a priori joint distribution of the $N_t N_r$ parameters $\beta_{k,\ell}(n)$, $\forall k, \ell$, and the multi-parameter BCRLB are very difficult to obtain. Therefore, using the assumption that phase noise of a practical oscillators varies slowly with time and the Taylor series approximation, we incorporate the phase noise innovations, $\Delta_\ell^{[t]}(n)$ and $\Delta_k^{[r]}(n)$, $\forall k, \ell$, in to the additive noise term, $w(n)$ and transform the joint estimation of channel gains, $\mathbf{A}$, and phase fluctuations, $\mathbf{B}(n)$ into a deterministic multi-parameter estimation problem over the length of the observation sequence.

III. CRAMÉR-RAO LOWER BOUNDS

In this section, new expressions for the Fisher's information matrices (FIMs) and the CRLBs for data-aided estimation (DAE) and decision-directed estimation (DDE) of phase noise and channel gains in MIMO systems are derived. Note that the derived CRLBs are applicable to *online* processing of the received signals for joint estimation of phase noise and channels.

A. CRLB for DAE

In order to coherently detect the transmitted signal at time instant n, $\mathbf{s}(n)$, the MIMO receiver needs to jointly estimate the channel gains and phase noise parameters, $\mathbf{A}$ and $\mathbf{B}(n)$, respectively. As a result, the vector parameter of interest $\boldsymbol{\lambda}$ is given by

$$\boldsymbol{\lambda} \triangleq \left[\boldsymbol{\lambda}_1^T, \ldots, \boldsymbol{\lambda}_{N_r}^T\right]^T \qquad (5)$$

where $\boldsymbol{\lambda}_k \triangleq \left[\boldsymbol{\alpha}_k^T, \boldsymbol{\beta}_k^T(n)\right]^T$, for $k = 1, \ldots, N_r$. Let us consider that a TS of length L_t is used to estimate the parameters of interest at time instant n, $\mathbf{A}$ and $\mathbf{B}(n)$. In the following steps, we seek to express the corresponding received signal vector as a function of the parameters of interest.

The phase noise model in (2) can be rewritten as

$$\theta_\ell^{[t]}(n - L_t) = \theta_\ell^{[t]}(n) - \sum_{m=n-L_t+1}^{n} \Delta_\ell^{[t]}(m)$$

$$\theta_k^{[r]}(n - L_t) = \theta_k^{[r]}(n) - \sum_{m=n-L_t+1}^{n} \Delta_k^{[r]}(m), \quad \forall k, \ell. \quad (6)$$

Using (6), the received signal at the kth receive antenna in (1) can be modelled as (7) at the bottom of the page.

For small values of L_t, (7b) can be tightly approximated as

$$y_k(n - L_t) \simeq \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)}$$
$$\times \left(1 - j\left(\sum_{m=n-L_t+1}^{n} \Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m)\right)\right)$$
$$\times s_\ell(n - L_t) + w_k(n - L_t), \qquad (8)$$

since for practical oscillators the phase noise innovation variances are small [11], [12] and the Taylor series expansion of the term, $e^{-j\left(\sum_{m=n-L_t+1}^{n}\left(\Delta_k^{[r]}(m)+\Delta_\ell^{[t]}(m)\right)\right)}$ for small phase innovations $\Delta_\ell^{[t]}$ and $\Delta_k^{[r]}$ can be approximated by

$$e^{-j\left(\sum_{m=n-L_t+1}^{n}\left(\Delta_k^{[r]}(m)+\Delta_\ell^{[t]}(m)\right)\right)}$$
$$\simeq 1 - j\left(\sum_{m=n-L_t+1}^{n}\left(\Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m)\right)\right). \quad (9)$$

Note that the small angle approximation in (9) has also been used in [7] and [57] for estimating and analyzing the effect of phase noise in SISO systems. Finally, Remark 5 at the end of this subsection compares the derived data-aided CRLB against the *posterior CRLB (PCRLB)* in [21] for SISO systems and shows that the above approximation is valid even for high phase noise variances, e.g., $\sigma_\Delta^2 = 10^{-3}$ rad^2 [11], [12], [19], [27], [29].

The received training signal at the kth antenna in (7) can be written in vector form as

$$\mathbf{y}_k \simeq \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell \underbrace{- j \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell \odot \boldsymbol{\psi}_{k,\ell} + \mathbf{w}_k}_{\text{overall noise}},$$
$$(10)$$

$$y_k(n - L_t) = \sum_{\ell=1}^{N_t} h_{k,\ell} e^{j\left(\theta_k^{[r]}(n) - \sum_{m=n-L_t+1}^{n} \Delta_k^{[r]}(m) + \theta_\ell^{[t]}(n) - \sum_{m=n-L_t+1}^{n} \Delta_\ell^{[t]}(m)\right)} s_\ell(n - L_t) + w_k(n - L_t) \qquad (7a)$$

$$= \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\left(\beta_{k,\ell}(n) - \sum_{m=n-L_t+1}^{n}\left(\Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m)\right)\right)} s_\ell(n - L_t) + w_k(n - L_t) \qquad (7b)$$

where

- $\mathbf{y}_k \triangleq [y_k(n-L_t+1), \ldots, y_k(n)]^T$;
- $\mathbf{s}_\ell \triangleq [s_\ell(n-L_t+1), \ldots, s_\ell(n)]^T$;
- $\mathbf{w}_k \triangleq [w_k(n-L_t+1), \ldots, w_k(n)]^T$; and
- $\boldsymbol{\psi}_{k,\ell} \triangleq \big[\sum_{m=n-L_t+2}^{n} \Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m), \sum_{m=n-L_t+3}^{n} \times \Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m), \ldots, 0 \big]^T$.

The received signals at all receive antennas, $\mathbf{y}$, can be written as $\mathbf{y} \triangleq [\mathbf{y}_1^T, \ldots, \mathbf{y}_{N_r}^T]^T$. Note that based on Assumption A2, $\Delta_k^{[r]}(n)$, $\Delta_{\bar{k}}^{[r]}(m)$, $\Delta_\ell^{[t]}(n)$, and $\Delta_{\bar{\ell}}^{[t]}(m)$, $\forall k \neq \bar{k}$, $\ell \neq \bar{\ell}$, $n \neq m$, are mutually independent. In addition, it follows from the assumptions in Section II and (1) that $w_k(n)$ and $w_{\bar{k}}(m)$, $\forall k \neq \bar{k}$, $n \neq \bar{m}$, and $\Delta_k^{[r]}(n)$, $\Delta_\ell^{[t]}(n)$, and $w_k(n)$, $\forall k, \ell, n$, are mutually independent. Subsequently, the received signal vector, $\mathbf{y}$, is distributed as $\mathcal{CN}(\boldsymbol{\mu}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{y}})$, where the mean and the $N_r L_t \times N_r L_t$ covariance matrix of $\mathbf{y}$, $\boldsymbol{\mu}_{\mathbf{y}}$ and $\boldsymbol{\Sigma}_{\mathbf{y}}$, respectively, are given by

$$\boldsymbol{\mu}_{\mathbf{y}} \triangleq \left[\boldsymbol{\mu}_{\mathbf{y}_1}^T, \ldots, \boldsymbol{\mu}_{\mathbf{y}_{N_r}}^T \right]^T, \tag{11a}$$

$$\boldsymbol{\Sigma}_{\mathbf{y}} \triangleq \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{y}_{1,1}} & \ldots & \boldsymbol{\Sigma}_{\mathbf{y}_{1,N_r}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{\mathbf{y}_{N_r,1}} & \ldots & \boldsymbol{\Sigma}_{\mathbf{y}_{N_r,N_r}} \end{bmatrix}. \tag{11b}$$

In (11), $\boldsymbol{\mu}_{\mathbf{y}_k}$ is given by $\boldsymbol{\mu}_{\mathbf{y}_k} = \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell$ and as shown in Appendix A, the $L_t \times L_t$ sub-matrices $\boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}$, for $\ell, \bar{\ell} = 1, \ldots, N_t$ and $k, \bar{k} = 1, \ldots, N_r$, can be determined as

$$\boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}} = \begin{cases} \sum_{\ell=1}^{N_t} \alpha_{k,\ell}^2 \mathbf{s}_\ell \mathbf{s}_\ell^H \odot (\sigma_{\Delta_k^{[r]}}^2 \boldsymbol{\Psi} + \sigma_{\Delta_\ell^{[t]}}^2 \boldsymbol{\Psi}) \\ \quad + \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1, \bar{\ell} \neq \ell}^{N_t} |\alpha_{k,\ell}||\alpha_{\bar{k},\bar{\ell}}| \\ \quad \times e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\bar{\ell}}(n)} \mathbf{s}_\ell \mathbf{s}_{\bar{\ell}}^H \odot (\sigma_{\Delta_k^{[r]}}^2 \boldsymbol{\Psi}) \\ \quad + \sigma_w^2 \mathbf{I}_{L_t \times L_t}, \qquad\qquad\qquad k = \bar{k} \\ \sum_{\ell=1}^{N_t} |\alpha_{k,\ell}||\alpha_{\bar{k},\bar{\ell}}| e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\bar{\ell}}(n)} \mathbf{s}_\ell \mathbf{s}_\ell^H \\ \quad \odot (\sigma_{\Delta_\ell^{[t]}}^2 \boldsymbol{\Psi}), \qquad\qquad\qquad k \neq \bar{k} \end{cases} \tag{12}$$

where $\boldsymbol{\Psi}$ is an $L_t \times L_t$ matrix given by

$$\boldsymbol{\Psi} \triangleq \begin{bmatrix} L_t-1 & L_t-2 & L_t-3 & \cdots & 0 \\ L_t-2 & L_t-2 & L_t-3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Proposition 1: The $2N_t N_r \times 2N_t N_r$ FIM for joint DAE of channel gains and phase noise parameters is given by

$$\mathbf{FIM} = \begin{bmatrix} \mathbf{FIM}_{1,1} & \ldots & \mathbf{FIM}_{1,N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{FIM}_{N_r,1} & \ldots & \mathbf{FIM}_{N_r,N_r} \end{bmatrix}. \tag{13}$$

In (13), $\mathbf{FIM}_{k,\bar{k}}$, for $k, \bar{k} = 1, \ldots, N_r$, are $2N_t \times 2N_t$ matrices that are determined as shown in (14) at the bottom of the page, where

- $\mathbf{U}_k \triangleq \left[\mathbf{0}_{(k-1)L_t \times N_t}^T, (\mathbf{E}_k \mathbf{S})^T, \mathbf{0}_{(N_r-k)L_t \times N_t}^T \right]^T$ is an $N_r L_t \times N_t$ matrix;
- $\mathbf{E}_k \triangleq \mathrm{diag}(e^{j\beta_{k,1}(n)}, \ldots, e^{j\beta_{k,N_t}(n)})$ is an $N_t \times N_t$ matrix, $\mathbf{S} \triangleq [\mathbf{s}_1, \ldots, \mathbf{s}_{N_t}]^T$ is an $N_t \times L_t$ matrix;
- $\boldsymbol{\Upsilon}_k \triangleq \mathrm{diag}(\alpha_{k,1}, \ldots, \alpha_{k,N_t})$ is an $N_t \times N_t$ matrix;
- the ℓth row and $\bar{\ell}$th column, for $\ell, \bar{\ell} = 1, \ldots, N_t$, elements of the $N_t \times N_t$ matrices $\boldsymbol{\Pi}_{11}$, $\boldsymbol{\Pi}_{12}$, $\boldsymbol{\Pi}_{21}$, and $\boldsymbol{\Pi}_{22}$ are given by

$$[\boldsymbol{\Pi}_{11}]_{\ell,\bar{\ell}}(k,\bar{k}) \triangleq \mathrm{Tr}\left\{ \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{k,\ell}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}} \right\},$$

$$[\boldsymbol{\Pi}_{12}]_{\ell,\bar{\ell}}(k,\bar{k}) \triangleq \mathrm{Tr}\left\{ \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{k,\ell}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}(n)} \right\},$$

$$[\boldsymbol{\Pi}_{21}]_{\ell,\bar{\ell}}(k,\bar{k}) \triangleq \mathrm{Tr}\left\{ \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{k,\ell}(n)} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}} \right\},$$

$$[\boldsymbol{\Pi}_{22}]_{\ell,\bar{\ell}}(k,\bar{k}) \triangleq \mathrm{Tr}\left\{ \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{k,\ell}(n)} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}(n)} \right\},$$

and the $N_r L_t \times N_r L_t$ matrices $\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{k,\ell}}$ and $\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{k,\ell}(n)}$ are derived in Appendix A.

Proof: See Appendix A. The following remarks are in order:

Remark 3: The CRLB for the estimation of channel gains and phase noise parameters, $\boldsymbol{\alpha}_k$ and $\boldsymbol{\beta}_k(n)$, for $k = 1, \ldots, N_r$, respectively, in (5) is given by

$$\mathbf{CRLB}(\boldsymbol{\lambda}) = \mathrm{diag}\left(\mathbf{FIM}^{-1}\right). \tag{15}$$

Given the structure of the FIM in (13) and (14), it is difficult to find a closed-form expression for the CRLB for an $N_r \times N_t$ MIMO system. However, partitioned matrix inverse [58] can be applied to find a closed-form expression for the CRLB for specific values of N_r and N_t.

Remark 4: The $N_t N_r \times N_t N_r$ FIM for estimation of overall channels in (4), $\mathbf{P}(n)$, $\mathbf{FIM}^{[\mathbf{P}(n)]}$ can be also determined in a similar fashion as above, where the $N_t \times N_t$ submatrices $\mathbf{FIM}_{k,\bar{k}}^{[\mathbf{P}(n)]}$, for $k, \bar{k} = 1, \ldots, N_r$, are given by

$$\mathbf{FIM}_{k,\bar{k}}^{[\mathbf{P}(n)]} = \left[2\Re\left\{ \left(\mathbf{U}_k^{[\mathbf{P}(n)]}\right)^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{U}_{\bar{k}}^{[\mathbf{P}(n)]} \right\} \right.$$
$$\left. + \boldsymbol{\Pi}_{11}^{[\mathbf{P}(n)]}(k,\bar{k}) \right]. \tag{16}$$

In (16), $\mathbf{U}_k^{[\mathbf{P}(n)]} \triangleq \left[\mathbf{0}_{(k-1)L_t \times N_t}^T, \mathbf{S}^T, \mathbf{0}_{(N_r-k)L_t \times N_t}^T \right]^T$, $\boldsymbol{\Pi}_{11}^{[\mathbf{P}(n)]}(k,\bar{k}) \triangleq \mathrm{Tr}\left\{ \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial p_{k,\ell}(n)} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial p_{\bar{k},\bar{\ell}}(n)} \right\}$, $p_{k,\ell}(n)$

$$\mathbf{FIM}_{k,\bar{k}} = \begin{bmatrix} 2\Re\left\{ \mathbf{U}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{U}_{\bar{k}} \right\} + \boldsymbol{\Pi}_{11}(k,\bar{k}) & -2\Im\left\{ \mathbf{U}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{U}_{\bar{k}} \boldsymbol{\Upsilon}_{\bar{k}} \right\} + \boldsymbol{\Pi}_{12}(k,\bar{k}) \\ 2\Im\left\{ \boldsymbol{\Upsilon}_k \mathbf{U}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{U}_{\bar{k}} \right\} + \boldsymbol{\Pi}_{21}(k,\bar{k}) & 2\Re\left\{ \boldsymbol{\Upsilon}_k \mathbf{U}_k^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \mathbf{U}_{\bar{k}} \boldsymbol{\Upsilon}_{\bar{k}} \right\} + \boldsymbol{\Pi}_{22}(k,\bar{k}) \end{bmatrix} \tag{14}$$
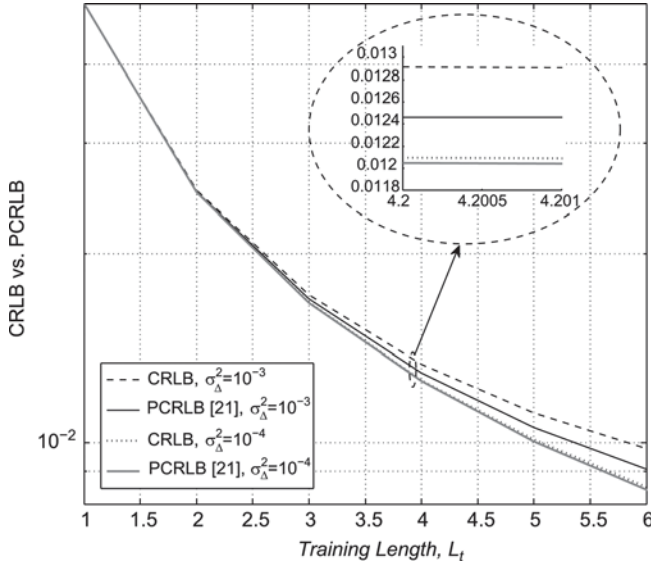
Fig. 3. The CRLB and PCRLB in [21] for data-aided estimation of phase noise in a SISO system with $\text{SNR} = 10$ dB and $h_{1,1} = 1$.

denotes the kth row and ℓth column element of $\mathbf{P}(n)$ and $\frac{\partial \mathbf{\Sigma_y}}{\partial p_{m,\ell}(n)}$, for $m = 1, \ldots, N_r$, is determined as

$$
\frac{\partial \mathbf{\Sigma}_{\mathbf{y}_{k,\bar{k}}}}{\partial p_{m,\ell}(n)} = \begin{cases} \begin{aligned} & p_{m,\ell}^*(n)\mathbf{s}_\ell \mathbf{s}_\ell^H \\ & \odot (\sigma^2_{\Delta_m^{[r]}} \mathbf{\Psi} + \sigma^2_{\Delta_\ell^{[t]}} \mathbf{\Psi}) \\ & + \sum_{\substack{p=1, \\ p \neq \ell}}^{N_t} \left(p_{m,p}^*(n)\mathbf{s}_\ell \mathbf{s}_p^H \right. \\ & \left. \odot \sigma^2_{\Delta_m^{[r]}} \mathbf{\Psi} \right), & m = k = \bar{k} \\ p_{\bar{k},\ell}(n)\mathbf{s}_\ell \mathbf{s}_\ell^H \odot \sigma^2_{\Delta_\ell^{[t]}} \mathbf{\Psi}, & m = k, \ \bar{k} \neq k \\ 0, & m = \bar{k}, \ k \neq \bar{k} \\ 0, & m \neq k, m \neq \bar{k}. \end{aligned} \end{cases}
$$

(17)

Using (15), the FIM in (17) can be applied to find the CRLB for estimation of the overall MIMO channels including phase noise, $\mathbf{P}(n)$. Note that as shown in Section V-B, $\mathbf{P}(n)$ is used at the MIMO receiver to equalize the effect of channels and phase noise. Therefore, the CRLB on the estimation accuracy of $\mathbf{P}(n)$ can be used to determine the effect of channel and phase noise estimation accuracy on the performance of MIMO systems.

Remark 5: Fig. 3 compares the data-aided CRLB for estimation of phase noise against the PCRLB in [21] for SISO systems. In addition to illustrating that the CRLB derived in this paper is accurate, Fig. 3 verifies the small angle approximation applied in this section for different values of L_t and phase noise variance in SISO systems. It is shown that even for large phase noise variances, e.g., $\sigma^2_\Delta = 10^{-3} \text{ rad}^2$, [11], [12], [19], [27], [29], the derived CRLB is close to the PCRLB in [21].

Remark 6: The FIM in (13) and (14) is not block diagonal. Therefore, the estimation of channel magnitudes and phase noise parameters in a MIMO system are coupled with one another, i.e., channel estimation accuracy is affected by the presence of phase noise and vice versa. This result indicates that channel and phase noise estimation need to be carried out jointly in a MIMO system.

Remark 7: The CRLB usually depends upon unknown parameters. Here, it is also intuitively reasonable that channel and phase noise estimation accuracies depend on their actual values, e.g., if the channel gains are extremely small, indicating low SNR conditions, the channel and phase noise estimation variances will be large [59].

B. CRLB for Decision-Directed Estimation (DDE)

Since in most practical scenarios of interest channel gains vary much slower than the phase noise processes [43] and based on the assumption of quasi-static fading channels, the estimates of the channel magnitudes obtained using the TSs and data-aided estimation can be used over the frame. Therefore, in the case of DDE, denote L_o as the number of previous symbols used to estimate the nth symbol's phase noise parameters, $\mathbf{B}(n)$, which consist of both pilots and estimates of data symbols. In order to make the analysis in this subsection tractable, we assume DDE with perfect decision feedback. Note that even though this assumption makes the proposed CRLB a looser bound for decision directed estimators with *imperfect* decision feedback, the numerical results in Section V demonstrate that the derived DDE-CRLB is a valid and accurate lower bound for both cases of perfect and imperfect decision feedback.

Based on the above set of assumptions the vector of parameters of interest for DDE, $\bar{\boldsymbol{\lambda}}$, is given by

$$
\bar{\boldsymbol{\lambda}} \triangleq \left[\boldsymbol{\beta}_1^T(n), \ldots, \boldsymbol{\beta}_{N_r}^T(n) \right]^T.
$$

(18)

The $N_t N_r \times N_t N_r$ FIM for the DDE, $\overline{\mathbf{FIM}}$, can be derived in a similar fashion as that of DAE in (13), where the $N_t \times N_t$ submatrices of $\overline{\mathbf{FIM}}$ denoted by $\overline{\mathbf{FIM}}_{k,\bar{k}}$, for $k, \bar{k} = 1, \ldots, N_r$, are given by[6]

$$
\overline{\mathbf{FIM}}_{k,\bar{k}} = 2\Re \left\{ \boldsymbol{\Upsilon}_k \bar{\mathbf{U}}_k^H \bar{\mathbf{\Sigma}}_{\bar{\mathbf{y}}}^{-1} \bar{\mathbf{U}}_{\bar{k}} \boldsymbol{\Upsilon}_{\bar{k}} + \bar{\mathbf{\Pi}}_{22}(k, \bar{k}) \right\}.
$$

(19)

In (19), $\bar{\mathbf{\Sigma}}_{\bar{\mathbf{y}}}$ denotes the covariance matrix of the observation sequence in the decision-directed scenario, $\bar{\mathbf{y}}_k \triangleq [y_k(n - L_o), \ldots, y_k(n-1)]^T$, which is determined by replacing $\mathbf{\Psi}$ with $\bar{\mathbf{\Psi}}$ in (12), where $\bar{\mathbf{\Psi}}$ is given by

$$
\bar{\mathbf{\Psi}} \triangleq \begin{bmatrix} L_o & L_o - 1 & L_o - 2 & \cdots & 1 \\ L_o - 1 & L_o - 1 & L_o - 2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}.
$$

In addition, in (19), $\bar{\mathbf{U}}_k \triangleq \left[\mathbf{0}_{(k-1)L_o \times N_t}^T, (\mathbf{E}_k \bar{\mathbf{S}})^T, \mathbf{0}_{(N_r-k)L_o \times N_t}^T \right]^T$ is an $N_r L_o \times N_t$ matrix, $\bar{\mathbf{S}} \triangleq [\bar{\mathbf{s}}_1, \ldots, \bar{\mathbf{s}}_{N_t}]^T$ is an $N_t \times L_o$ matrix, $\bar{\mathbf{s}}_\ell \triangleq [\bar{s}_\ell(n - L_o), \ldots, \bar{s}_\ell(n-1)]^T$, $\mathbf{E}_k$ and $\boldsymbol{\Upsilon}_k$ are defined below (14), and $\bar{\mathbf{\Pi}}_{22}(k, \bar{k})$ is determined by replacing $\mathbf{\Sigma_y}$ with $\bar{\mathbf{\Sigma}}_{\bar{\mathbf{y}}}$ in $\mathbf{\Pi}_{22}(k, \bar{k})$ below (14). The CRLB for the decision directed case, $\overline{\mathbf{CRLB}}$ is given by

$$
\overline{\mathbf{CRLB}}(\bar{\boldsymbol{\lambda}}) = \text{diag}\left(\overline{\mathbf{FIM}}^{-1} \right).
$$

[6]Similar steps as the ones outlined in Section III-A and Appendix A are used and, therefore, are not repeated here.

IV. CHANNEL AND PHASE NOISE ESTIMATION

In this section, an LS estimator for joint data-aided estimation of channels and phase noise parameters and a WLS estimator and EKF for decision-directed tracking of phase noise parameters over a frame are derived.

A. Data-Aided Estimation (DAE)

The conditional likelihood function of $\mathbf{y}$ given the channel magnitudes and phase noise parameters at time instant n, $\mathbf{A}$ and $\mathbf{B}(n)$, respectively, is given by

$$p\left(\mathbf{y}\big|\mathbf{A}, \mathbf{B}(n)\right) = \frac{1}{(\pi)^{N_r L_t} \det(\boldsymbol{\Sigma}_{\mathbf{y}})}$$
$$\times \exp\left\{-(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^H \boldsymbol{\Sigma}_{\mathbf{y}}^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})\right\}. \quad (20)$$

Based on (20) and since the covariance matrix, $\boldsymbol{\Sigma}_{\mathbf{y}}$ in (11) and (12), is a function of both channel magnitudes and phase noises, the derivation of the joint maximum-likelihood estimator of channel and phase noise parameters is highly complex. Therefore, in this section we derive a significantly less complex data-aided LS estimator. Using (20), the joint LS estimates of $\mathbf{A}$ and $\mathbf{B}(n)$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}^{[\text{DAE}]}$, respectively, can be determined as

$$\hat{\mathbf{A}}, \hat{\mathbf{B}}^{[\text{DAE}]}(n) = \arg\min_{\mathbf{A},\mathbf{B}}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^H(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}). \quad (21)$$

The cost function in (21) can be modified as

$$\hat{\mathbf{A}}, \hat{\mathbf{B}}^{[\text{DAE}]}(n) = \arg\min_{\mathbf{A},\mathbf{B}}(\mathbf{Y} - \mathbf{P}(n)\mathbf{S})^H(\mathbf{Y} - \mathbf{P}(n)\mathbf{S}) \quad (22)$$

where $\mathbf{Y} \triangleq [\mathbf{y}_1, \ldots, \mathbf{y}_{N_r}]^T$ and $\mathbf{S}$ and $\mathbf{P}(n)$ are defined in (14) and (4), respectively. Using well-known methods, the LS estimate of $\mathbf{P}(n)$, $\hat{\mathbf{P}}(n)$, can be shown to be

$$\hat{\mathbf{P}}(n) = \mathbf{Y}\mathbf{S}^H(\mathbf{S}\mathbf{S}^H)^{-1} = \frac{\mathbf{Y}\mathbf{S}^H}{L_t} \quad (23)$$

where the second equality in (23) follows from Assumption A1, i.e., $\mathbf{S}\mathbf{S}^H = L_t \mathbf{I}_{N_t \times N_t}$. As demonstrated in [11], [12] and discussed in Section II, for most practical oscillators, the phase noise variances, $\sigma^2_{\Delta^{[t]}_\ell}$ and $\sigma^2_{\Delta^{[r]}_k}$, $\forall k, \ell$ in (3), are very small. Therefore, the LS estimator is expected to accurately obtain the MIMO channel and phase noise parameters. Using (23), estimates of the channel magnitude and phase noise matrices are determined by

$$\hat{\mathbf{A}} = \frac{\left|\mathbf{Y}\mathbf{S}^H\right|}{L_t} \quad \text{and} \quad \hat{\mathbf{B}}^{[\text{DAE}]}(n) = \angle\left\{\frac{\mathbf{Y}\mathbf{S}^H}{L_t}\right\}, \quad (24)$$

respectively. Simulation results in Section V-B show that the performance of the proposed LS estimator in (24) is close to the CRLB over a wide range of SNR values.

B. Decision-Directed Estimation

In order to track the phase noise parameters throughout a frame, in this subsection WLS and EKF decision-directed estimators are proposed.

1) Weighted Least-squares Estimator: The conditional *log-likelihood function (LLF)* up to an additive constant can be determined as

$$\log\left(p(\bar{\mathbf{y}}|\hat{\mathbf{A}}, \mathbf{B}(n))\right)$$
$$= \log\left(\det(\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}})\right) + (\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})^H \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}}), \quad (25)$$

where $\bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}_k} = \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \bar{\mathbf{s}}_\ell$, and $\bar{\mathbf{y}}, \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}, \bar{\mathbf{s}}_\ell$ are defined in Section III-B. Using (25), decision-directed WLS estimates of the nth symbol's $N_t N_r$ phase noises, $\hat{\mathbf{B}}^{[\text{DDE}]}(n)$, is given by

$$\hat{\mathbf{B}}^{[\text{DDE}]}(n) = \arg\min_{\mathbf{B}(n)}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})^H \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1}(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}}). \quad (26)$$

In order to reduce the computational complexity, the term $\log\left(\det(\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}})\right)$ in (26) is omitted, resulting in a WLS estimator. Numerical results demonstrate that $\log\left(\det(\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}})\right)$ in (12) does not vary significantly for different phase noise values, $\beta_{k,\ell}, \forall k, \ell$.

Note that the minimization in (26) requires an exhaustive search over an $N_t N_r$ dimensional space, which is computationally intensive. Even though the proposed WLS estimator may be computationally complex to implement, it is used to verify the CRLB derivations and to assess the performance of the proposed EKF estimator. In addition, the computational complexity of this exhaustive search and the proposed WLS estimator can be reduced by applying *alternating projection (AP)* [60]. AP is an iterative process, where at each iteration the right hand side of (26) is minimized with respect to one of the phase noise parameters, e.g., $\beta_{k,\ell}$, while the remaining terms are kept at their most updated values. In other words, AP reduces the multi-dimensional minimization problem into a series of one-dimensional minimizations. Even though AP is not guaranteed to converge to the true estimates, in [60, Sec. IV-A] it is shown that AP converges to a local maximum and through proper initialization, it results in global convergence. Since phase noise does not vary quickly with time, the phase noise estimates obtained at the beginning of each frame, using the data-aided LS estimator in Section IV-A, are used to initialize the AP process for the WLS estimator. This ensures that the iterative algorithm converges more quickly. In addition, the numerical simulations in Fig. 6 indicate that using the above initialization, AP converges to the *true* estimates in only 4 cycles.

2) Extended Kalman Estimator: This section presents an EKF to track $N_r N_t$ phase noise parameters $\beta_{k,\ell}(n)$, $\forall k, \ell$, in decision-directed mode. First, the *state* and *observation* are developed for the EKF. Using (2), $\beta_{k,\ell}(n)$, $\forall k, \ell$, can be written as

$$\beta_{k,\ell}(n) = \beta_{k,\ell}(n-1) + \Delta_{k,\ell}(n) \quad (27)$$

where $\Delta_{k,\ell}(n) = \Delta^{[r]}_k(n) + \Delta^{[t]}_\ell(n)$ is the sum of the phase noise innovations for the ℓth transmit and the kth receive antennas. Using (27), the unknown *state vector* $\boldsymbol{\phi}(n) \triangleq [\boldsymbol{\beta}_1^T(n), \ldots, \boldsymbol{\beta}_{N_r}^T(n)]^T$ is given by

$$\boldsymbol{\phi}(n) = \boldsymbol{\phi}(n-1) + \boldsymbol{\delta}(n) \quad (28)$$

where the state noise vector

$$\boldsymbol{\delta}(n)$$
$$\triangleq \left[\Delta_{1,1}(n), \ldots, \Delta_{1,N_t}(n), \ldots, \Delta_{N_r,1}(n), \ldots, \Delta_{N_r,N_t}(n)\right]^T$$

is distributed as $\boldsymbol{\delta}(n) \sim \mathcal{N}(\mathbf{0}_{N_t N_r \times 1}, \mathbf{Q})$. The state noise co-variance matrix, $\mathbf{Q} = \mathbb{E}\left\{\boldsymbol{\delta}(n)\boldsymbol{\delta}(n)^T\right\}$, is an $N_t N_r \times N_t N_r$ matrix that can be determined straightforwardly as

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{1,1} & \cdots & \mathbf{Q}_{1,N_r} \\ \vdots & \ddots & \vdots \\ \mathbf{Q}_{N_r,1} & \cdots & \mathbf{Q}_{N_r,N_r} \end{bmatrix}, \qquad (29)$$

$$\mathbf{Q}_{k,\bar{k}} = \begin{cases} \sigma^2_{\Delta_k^{[r]}} \mathbf{1}_{N_t \times N_t} + \mathrm{diag}\left(\sigma^2_{\Delta_1^{[t]}}, \ldots, \sigma^2_{\Delta_{N_t}^{[t]}}\right), & k = \bar{k} \\ \mathrm{diag}\left(\sigma^2_{\Delta_1^{[t]}}, \ldots, \sigma^2_{\Delta_{N_t}^{[t]}}\right), & k \neq \bar{k}. \end{cases}$$
$$(30)$$

The *observation* equation for the Kalman filter is given by

$$\mathbf{y}(n) = \mathbf{z}(n) + \mathbf{w}(n)$$
$$= \left(\mathbf{A} \odot e^{j\mathbf{B}(n)}\right)\mathbf{s}(n) + \mathbf{w}(n), \qquad (31)$$

where $\mathbf{z}(n) \triangleq [z_1(n), \ldots, z_{N_r}(n)]^T = \left(\mathbf{A} \odot e^{j\mathbf{B}(n)}\right)\mathbf{s}(n)$. Since the observation equation in (31) is a non-linear function of the unknown state vector $\boldsymbol{\phi}(n)$, the EKF is used instead. The EKF uses Taylor series expansion to linearize the non-linear observation equation in (31) about the current estimates [61]. Thus, the Jacobian matrix is evaluated by computing the first order partial derivative of $\mathbf{z}(n)$ with respect to the state vector $\boldsymbol{\phi}(n)$ as

$$\mathbf{D}\big(\boldsymbol{\phi}(n)\big) = \frac{\partial \mathbf{z}(n)}{\partial \boldsymbol{\phi}^T(n)} = \begin{bmatrix} \frac{\partial z_1(n)}{\partial \boldsymbol{\beta}_1^T(n)} & \mathbf{0}_{1 \times N_t} & \cdots & \mathbf{0}_{1 \times N_t} \\ \mathbf{0}_{1 \times N_t} & \frac{\partial z_2(n)}{\partial \boldsymbol{\beta}_2^T(n)} & \cdots & \mathbf{0}_{1 \times N_t} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times N_t} & \mathbf{0}_{1 \times N_t} & \cdots & \frac{\partial z_{N_r}(n)}{\partial \boldsymbol{\beta}_{N_r}^T(n)} \end{bmatrix},$$
$$(32)$$

where $\mathbf{D}\big(\boldsymbol{\phi}(n)\big)$ denotes the $N_r \times N_r N_t$ Jacobian matrix evaluated at $\boldsymbol{\phi}(n)$, and

$$\frac{\partial z_k(n)}{\partial \boldsymbol{\beta}_k^T(n)} = \left[\frac{\partial z_k(n)}{\partial \beta_{k,1}(n)}, \ldots, \frac{\partial z_k(n)}{\partial \beta_{k,N_t}(n)}\right]$$
$$= \left[js_1(n)\alpha_{k,1}e^{j\beta_{k,1}(n)}, \ldots, \right.$$
$$\left. js_{N_t}(n)\alpha_{k,N_t}e^{j\beta_{k,N_t}(n)}\right]. \qquad (33)$$

Using (28)–(33), the remaining EKF equations can be formulated as (34)–(38) at the bottom of the page, where

- $\mathbf{D}\big(\boldsymbol{\phi}(n|n-1)\big) = \mathbf{D}\big(\boldsymbol{\phi}(n)\big)\big|_{\boldsymbol{\phi}(n)=\hat{\boldsymbol{\phi}}(n|n-1)}$;
- $\hat{\boldsymbol{\phi}}(n|n-1) = \left[\hat{\boldsymbol{\beta}}_1^T(n|n-1), \ldots, \hat{\boldsymbol{\beta}}_{N_r}^T(n|n-1)\right]^T$ is the predicted state vector at the nth symbol;
- $\mathbf{W} = \sigma_w^2 \mathbf{I}_{N_r \times N_r}$;
- $\hat{\mathbf{B}}(n) = \mathbf{B}(n)\big|_{\boldsymbol{\phi}(n)=\hat{\boldsymbol{\phi}}(n|n-1)}$;
- $\mathbf{K}(n)$ is the $N_t N_r \times N_r$ Kalman gain matrix;
- $\hat{\mathbf{A}}$ is given in (24); and
- $\mathbf{M}(n|n)$ is the $N_t N_r \times N_t N_r$ filtering error covariance matrix.

Note that the filtered state vector estimate, $\hat{\boldsymbol{\phi}}(n|n)$, and the $N_t N_r \times N_t N_r$ filtering error covariance matrix, $\mathbf{M}(n|n)$, are updated every nth symbol using the Kalman gain $\mathbf{K}(n)$. Before starting the EKF recursion (32)–(38), $\hat{\boldsymbol{\phi}}(L_t|L_t)$ and $\mathbf{M}(L_t|L_t)$ should be initialized with appropriate values, where L_t corresponds to the last training symbol. It is assumed that the Kalman filter initializes the state vector as $\hat{\boldsymbol{\phi}}(L_t|L_t) = \mathrm{vec}\left(\left(\hat{\mathbf{B}}^{[\mathrm{DAE}]}\right)^T(L_t)\right)$, where $\left(\hat{\mathbf{B}}^{[\mathrm{DAE}]}\right)(L_t)$ is found using (24). The error covariance matrix, $\mathbf{M}(L_t|L_t)$, is initialized as $\mathbf{M}(L_t|L_t) = \chi \mathbf{I}_{N_r N_t \times N_r N_t}$, where χ is a constant that is used to adjust the reliance of the EKF on the data-aided estimates obtained using the LS estimator[7].

C. Complexity Analysis

In this subsection, the computational complexity of the WLS and EKF in decision-directed mode is analyzed. Throughout this section, computational complexity is defined as the number of complex additions plus number of multiplications [59] required to update the phase noise estimates at every symbol. Let us denote the computational complexity of the WLS algorithm

[7]In this paper χ is generally selected to be a small value, e.g, $\chi = 0.1$, since as shown in Section V-B the data-aided estimates are accurate.

$$\hat{\boldsymbol{\phi}}(n|n-1) = \hat{\boldsymbol{\phi}}(n-1|n-1), \qquad (34)$$

$$\mathbf{M}(n|n-1) = \mathbf{M}(n-1|n-1) + \mathbf{Q}, \qquad (35)$$

$$\mathbf{K}(n) = \mathbf{M}(n|n-1)\mathbf{D}^H\big(\boldsymbol{\phi}(n|n-1)\big)\left(\mathbf{D}\big(\boldsymbol{\phi}(n|n-1)\big)\mathbf{M}(n|n-1)\mathbf{D}^H\big(\boldsymbol{\phi}(n|n-1)\big) + \mathbf{W}\right)^{-1}, \qquad (36)$$

$$\hat{\boldsymbol{\phi}}(n|n) = \hat{\boldsymbol{\phi}}(n|n-1) + \Re\left\{\mathbf{K}(n)\left[\mathbf{y}(n-1) - \left(\hat{\mathbf{A}} \odot e^{j\hat{\mathbf{B}}(n)}\right)\mathbf{s}(n-1)\right]\right\}, \qquad (37)$$

$$\mathbf{M}(n|n) = \Re\{\mathbf{M}(n|n-1) - \mathbf{K}(n)\mathbf{D}\big(\boldsymbol{\phi}(n|n-1)\big)\mathbf{M}(n|n-1)\}, \qquad (38)$$

by $C_{\mathrm{WLS}} = C_{\mathrm{WLS}}^{[M]} + C_{\mathrm{WLS}}^{[A]}$. The terms $C_{\mathrm{WLS}}^{[M]}$ and $C_{\mathrm{WLS}}^{[A]}$ denote the number of complex multiplications and additions, respectively, used by the WLS estimator and are determined as shown in (39) and (40) at the bottom of the page, where $\mathcal{N}$ denotes the number of alternating projection cycles used, and κ denotes the step size in the WLS search in (26). Similarly, the computational complexity of the EKF algorithm is denoted by $C_{\mathrm{KAL}} = C_{\mathrm{KAL}}^{[M]} + C_{\mathrm{KAL}}^{[A]}$. The notations $C_{\mathrm{KAL}}^{[M]}$ and $C_{\mathrm{KAL}}^{[A]}$ are used to denote the number of complex multiplications and additions, respectively, used by the EKF estimator and are determined as

$$
C_{\mathrm{KAL}}^{[M]} = \underbrace{2N_t N_r^2 (N_t N_r + N_r) + N_r^3}_{(36)}
$$
$$
+ \underbrace{N_r N_t (2 + N_r)}_{(37)} + \underbrace{N_r^3 N_t^2 (1 + N_t)}_{(38)} + \underbrace{2N_r N_t}_{(32)},
$$
(41)

$$
C_{\mathrm{KAL}}^{[A]} = \underbrace{(N_r N_t)^2}_{(35)} + \underbrace{N_r (N_r N_t - 1)(2N_r N_t + N_r)}_{(36)}
$$
$$
+ \underbrace{N_r^2 [1 + N_t (N_r - 1)] + N_r^3}_{(36)}
$$
$$
+ \underbrace{N_r [N_t (N_r - 1) + 2N_t]}_{(37)}
$$
$$
+ \underbrace{(N_r N_t)^2 (N_t N_r + N_r - 1)}_{(38)}.
$$
(42)

Remark 8: In order to quantitatively compare the computational complexity of the proposed WLS and EKF estimator, we evaluate C_{WLS} and C_{KAL} for a 2×2 MIMO system. We assume $L_o = N_t$ and in order to ensure that the performance of the proposed WLS is close to the CRLB, the step size, $\kappa = 10^{-2}$ and $\mathcal{N} = 4$ in (39) and (40), respectively. It is observed that the proposed EKF estimator is 3070 times more computationally efficient than the proposed WLS estimator.

V. NUMERICAL RESULTS AND DISCUSSIONS

In this section, we evaluate the performance of the proposed estimators versus the CRLB. Subsequently, the BER of a MIMO system employing the proposed channel and phase noise estimators is investigated in detail. Throughout this section $L_t = L_o = N_t$, $\chi = 0.1$, and without loss of generality, it is assumed that $\sigma_{\Delta_\ell^{[t]}}^2 = \sigma_{\Delta_k^{[r]}}^2 = \sigma_\Delta^2$, $\forall k, \ell$, and $\sigma_w^2 = \frac{1}{\mathrm{SNR}}$. The MIMO channel matrix is generated as a sum of LoS and *non-line-of-sight (NLoS)* components such that the overall channel matrix, $\mathbf{H}$, is given by [14]

$$
\mathbf{H} = \sqrt{\frac{K}{K+1}} \mathbf{H}_{\mathrm{LoS}} + \sqrt{\frac{1}{K+1}} \mathbf{H}_{\mathrm{NLoS}} \qquad (43)
$$

where K denotes the Rician factor, i.e., the ratio between the power of the LoS and NLoS channel components [62, p. 52], $\mathbf{H}_{\mathrm{LoS}}$ and $\mathbf{H}_{\mathrm{NLoS}}$, respectively[8]. The elements of $\mathbf{H}_{\mathrm{LoS}}$ are generated according to the model in [14] with the antenna spacing at the transmitters and receiver, d_t and d_r, respectively, set to their optimum values, the distance between transmitter and receiver is set to 2000 m, the antenna angles at the transmitter and receiver, φ_t and φ_r, respectively, are set to $\varphi_t = 5°$ and $\varphi_r = 0°$, the carrier frequency is set to 60 GHz, and unless otherwise specified $K = 2$ dB. The elements of $\mathbf{H}_{\mathrm{NLoS}}$ are modeled asindependent and identically distributed (i.i.d) complex Gaussian random variables with $\mathcal{CN}(0, 1)$. Walsh-Hadamard codes with binary phase-shift keying (BPSK) or quadrature phase-shift keying (QPSK) are used for the TSs. Given that the estimation range of the proposed WLS and EKF estimators are limited to $[-\pi, \pi)$ the phase unwrapping algorithm in [3] is applied here, where phase noise estimates for prior symbols are used in combination with the phase noise variance to unwrap the estimate for the current symbol. A minimum of 10^5 Monte Carlo trials are used. Finally, the mean-square error (MSE) performance of the proposed estimators and the BER performance of the overall MIMO system are investigated in the following subsections.

A. Estimation Performance

Channel realizations are drawn independently from the model in (43) for each Monte Carlo simulation trial. BPSK modulation is used for the pilot and data symbols. Without loss of generality, only the MSE for the estimation of channel gain and phase noise

---

[8]Note that $K = -\infty$ and $K = \infty$ dB result in pure NLoS and LoS channels, respectively [14].

$$
C_{\mathrm{WLS}}^{[M]} = \mathcal{N}(N_t N_r) \frac{2\pi}{\kappa} \Bigg\{ \underbrace{N_r \left[N_t (2L_o^2) + N_t (N_t - 1)(2L_o^2 + 3) \right] + (N_r^2 - N_r) \left[N_t (2L_o^2 + 3) \right]}_{\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}} \text{ in } (26)}
$$
$$
+ \underbrace{(N_r L_o)^3}_{(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})^H \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1} (\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})} + \underbrace{(N_r L_o)^3}_{\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1} \text{ in } (26)} + \underbrace{N_r N_t (1 + L_o)}_{\bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}} \text{ in } (26)} \Bigg\}, \qquad (39)
$$

$$
C_{\mathrm{WLS}}^{[A]} = \mathcal{N}(N_t N_r) \frac{2\pi}{\kappa} \Bigg\{ \underbrace{N_r \left[L_o^2 \{ (N_t - 1)^2 + N_t \} \right] + (N_r^2 - N_r) \left[L_o^2 (N_t - 1) \right]}_{\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}} \text{ in } (26)} + \underbrace{N_r L_o (N_r L_o - 1)^2}_{(\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})^H \bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1} (\bar{\mathbf{y}} - \bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}})}
$$
$$
+ \underbrace{(N_r L_o)^3}_{\bar{\boldsymbol{\Sigma}}_{\bar{\mathbf{y}}}^{-1} \text{ in } (26)} + \underbrace{N_r L_o (N_t - 1)}_{\bar{\boldsymbol{\mu}}_{\bar{\mathbf{y}}} \text{ in } (26)} \Bigg\} \qquad (40)
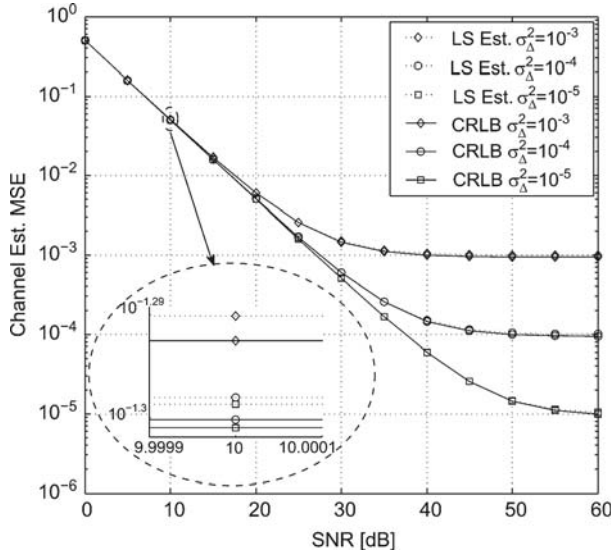$$

Fig. 4. MSE for DAE of channel, $\mathbf{P}(n)$, for a 2×2 MIMO system for different phase noise variances.

Fig. 5. MSE for DAE of phase noise for a 2×2 MIMO system for different phase noise variances.

for the first antenna element is presented. Note that similar results are obtained for the estimation of the parameters for the remaining antennas and are not presented here to avoid repetition.

Figs. 4 and 5 plot the CRLB and MSE for DAE of MIMO channels, $\mathbf{P}(n)$, and time-varying phases, respectively, versus SNR. The CRLB in (15) is numerically evaluated for different phase noise variances, e.g., $\sigma_\Delta^2 = [10^{-3}, 10^{-4}, 10^{-5}] \text{ rad}^2$. Note that, $\sigma_\Delta^2 = 10^{-3} \text{ rad}^2$, corresponds to a very high phase noise variance [11], [12], [19], [27], [29]. The CRLB results in Fig. 4 show that in the presence of phase noise, estimation of the MIMO channel suffers from an error floor, which is directly related to the variance of phase noise innovations. This result can be anticipated given that as shown in (10) even at very high SNR, where the effect of the AWGN noise, $\mathbf{w}_k$, is negligible, the additive noise corresponding to phase noise limits the estimation accuracy. The same noise also limits the estimation accuracy of the nth symbol's phase noise parameters in the data-aided case and results in an error floor in Fig. 5. Using Monte Carlo simulations the MSE of the data-aided LS estimator for jointly estimating the MIMO channels and phases in (24) is also evaluated and compared to the CRLB. The results in Figs. 4 and 5 show that the proposed LS estimator's MSE for channel estimation is close to the CRLB over a wide range of SNR values while its MSE for phase estimation is slightly higher than the CRLB at low SNR.

Fig. 6 plots the CRLB and MSE of the proposed WLS and EKF for DDE of phase noise versus SNR. The CRLB for DDE with perfect decision feedback in (19) is numerically evaluated for different phase noise variances and $L_o = 2$. As depicted in Fig. 6, the CRLB for DDE of phase noise also suffers from an error floor, which is higher compared to the DAE in Fig. 5. This higher error floor is due to the Wiener model of phase noise in (2) and since in the case of DDE the transmitted symbols at time instant n are assumed to be unknown, i.e., only the observations up to the $(n-1)$th symbol, $\bar{\mathbf{y}}_k \triangleq [y_k(n-L_o), \ldots, y_k(n-1)]^T$ can be used while estimating the
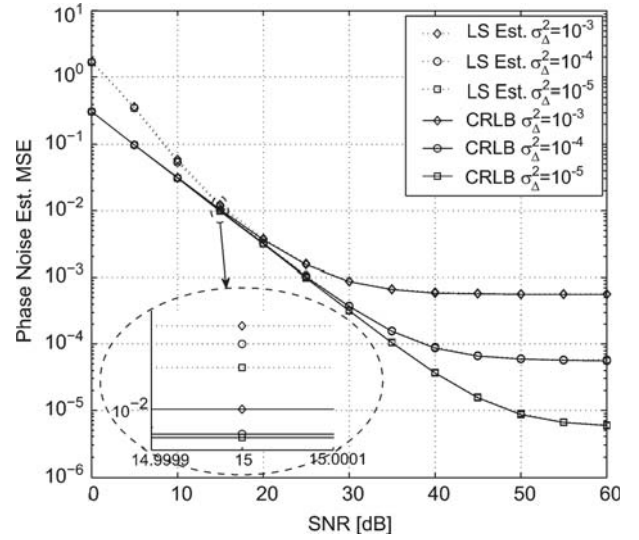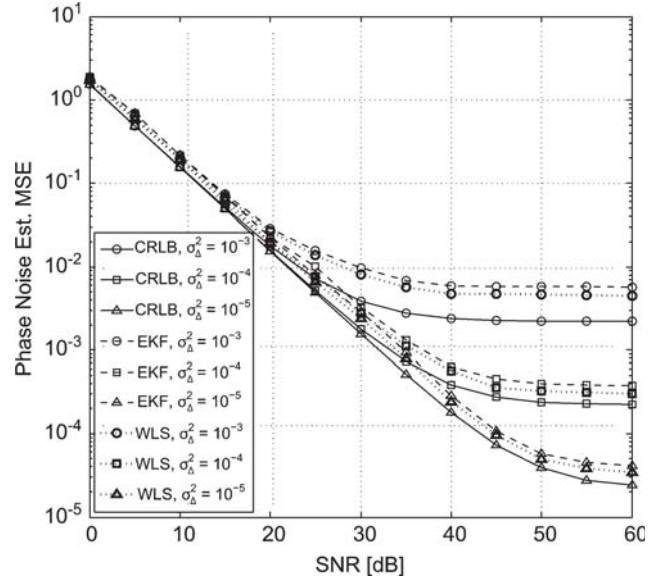

Fig. 6. Perfect decision feedback MSE for DDE of phase noise for a 2×2 MIMO system for different phase noise variances.

nth symbol's phase noise. Therefore, as observed in Fig. 6 the estimation accuracy of phases at time n, $\mathbf{B}(n)$, cannot reach below the variance of the white Gaussian phase noise innovations, $\Delta_\ell^{[t]}(n)$ and $\Delta_k^{[r]}(n)$, $\forall k, \ell$.

In Fig. 6, the MSE performances of the proposed WLS and EKF estimators are also compared against the CRLB, where it is shown that the proposed estimators' MSEs are close to the CRLB over a wide range of SNR values. In order to ensure a fair comparison with the CRLB, the MSE performances of the proposed WLS in (26) and EKF estimators in (32)–(38) are also evaluated with perfect decision feedback. Note that even though the Kalman filter is an optimal *minimum mean-square error (MMSE)* estimator [55], the extended Kalman filter does not have the same optimality properties and its performance highly depends on the accuracy of the applied linearization [55]. Therefore, as shown in Fig. 6, as the phase noise variances increase
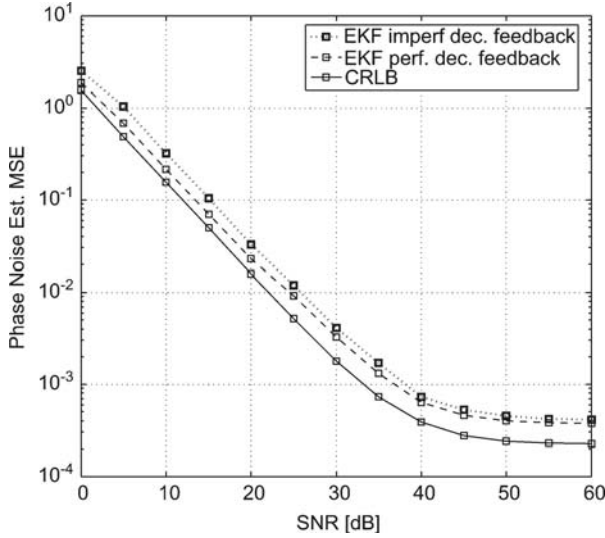
Fig. 7. Imperfect vs. perfect decision feedback MSE for DDE of phase noise for a 2×2 MIMO system ($\sigma_\Delta^2 = 10^{-4}\ \mathrm{rad}^2$ and $L_p = 10$).

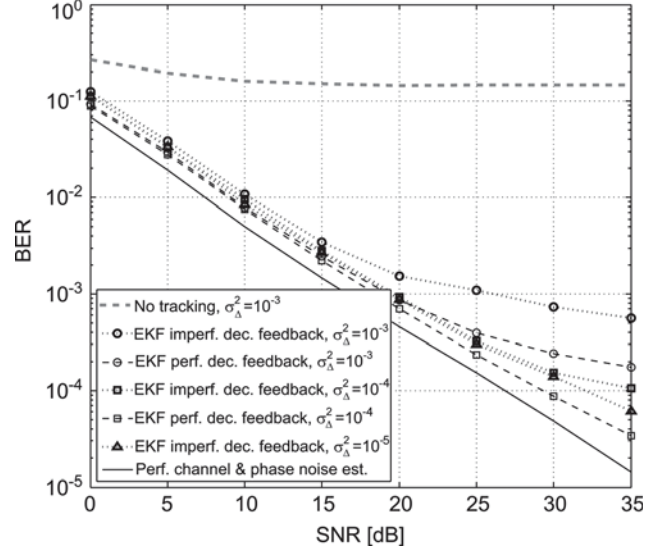Fig. 8. Perfect and Imperfect decision feedback BER of a 2×2 MIMO system for phase noise variance, $\sigma_\Delta^2 = [10^{-3}, 10^{-4}, 10^{-5}]\ \mathrm{rad}^2$, (BPSK and $L_p = 10$).

and the accuracy of the applied linearization for EKF decreases, the gap between the MSE of the proposed EKF and the CRLB slightly widens.

Fig. 7 compares the MSE of the proposed EKF estimator for both perfect and imperfect decision feedback against the CRLB in (19). It is shown that the proposed EKF estimator's performance while operating in imperfect decision feedback mode is close to that of perfect decision feedback and the CRLB in (19) over a wide range of SNRs. The latter indicates that even though the CRLB in (19) is derived based on the assumption of perfect decision feedback, it can be used as a tight bound for assessing the performance of imperfect decision feedback based estimators at medium-to-high SNR values. Finally, note that since the estimator in [13] fails to accurately track phase noise with the proposed system setup, its performance is not depicted in this subsection.

B. MIMO System Performance

In this section the BER performance of uncoded 2×2 and 4×4 MIMO systems in the presence of phase noise is investigated. The proposed LS estimator and training symbols at the start of each frame are used to obtain the MIMO channel gains and phases at time n. These estimates are used to initialize the state model of the one step ahead EKF estimator and obtain the phase noise estimates for the data symbols at time $n + 1$. Subsequently, the detected data symbols at time $n + 1$ are used to estimate the phase noise processes at time $n + 2$ and this process is carried out throughout the frame to the last symbol. The frame length is set to $L_f = 1000$ symbols and new channels are generated for each frame. Unless otherwise specified, the pilot spacing is set to $L_p = 10$, which corresponds to a synchronization overhead of 10%. An MMSE linear receiver given by

$$\hat{\mathbf{s}}(n) = \left(\hat{\mathbf{P}}^H(n)\hat{\mathbf{P}}(n) + \sigma_w^2 \mathbf{I}_{N_t \times N_t}\right)^{-1} \hat{\mathbf{P}}^H(n)\mathbf{y}(n) \quad (44)$$

is used to equalize the effect of phase noise and channel gains. Note that since the matrices $\mathbf{\Theta}^{[r]}(n)$ and $\mathbf{\Theta}^{[t]}(n)$ in (4) are diagonal, $\forall n$, phase noise does not affect the conditioning of the overall MIMO channel matrix. Finally, the BER of a MIMO system using the proposed decision-directed WLS estimator is not presented, given that its application in practical settings is limited due to its extremely high complexity as shown in Section IV-C.

Fig. 8 depicts the BER performance of a 2×2 MIMO system using the proposed LS channel and phase estimator *without* phase tracking and with the proposed EKF phase tracking in the presence of various phase noise variances. The scenario with perfect channel estimation and synchronization is also plotted and is used as a benchmark for assessing the performance of the proposed channel and phase tracking system. The results in Fig. 8 demonstrate that without phase tracking throughout the frame, the MIMO system performance deteriorates significantly. On the other hand, by combining the proposed LS and EKF channel and phase estimators, respectively, the BER performance of the MIMO system is shown to improve immensely even in the presence of very strong phase noise, e.g., $\sigma_\Delta^2 = 10^{-3}\ \mathrm{rad}^2$. In addition, it is demonstrated that the proposed EKF is capable of tracking the phase noise accurately with imperfect decision feedback where the gap between the perfect and imperfect decision feedback scenarios is small (a performance gap of 1.5 dB with $\mathrm{SNR} = 20$ dB and $\sigma_\Delta^2 = 10^{-4}\ \mathrm{rad}^2$). More importantly, it is shown that the BER performance of a MIMO system using the combination of the proposed channel and phase noise estimators is close to the ideal case of perfect channel and phase noise estimation (a performance gap of 3 dB with SNR=20 dB). Finally, Fig. 8 shows that the overall MIMO system's BER performance suffers from an error floor at high SNR. This result is anticipated, since at high SNR the performance of the MIMO system is dominated by phase noise instead of AWGN and as depicted in Fig. 6, the effect of phase noise cannot be completely eradicated.
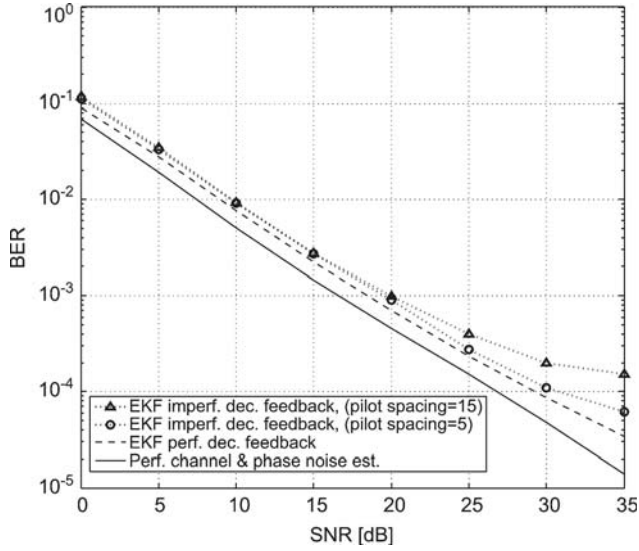
Fig. 9. Perfect and Imperfect decision feedback BER of a 2×2 MIMO system for different pilot spacing $L_p = [5, 15]$ (BPSK and $\sigma_\Delta^2 = 10^{-4} \, \mathrm{rad}^2$).

Fig. 10. BER of a 2×2 MIMO system for QPSK and 16-QAM modulations ($L_p = 10$ and $\sigma_\Delta^2 = 10^{-4} \, \mathrm{rad}^2$).

Fig. 9 illustrates the BER performance of a 2×2 MIMO system using the proposed EKF phase tracking algorithm with imperfect decision feedback for different pilot spacing. Based on the results in Fig. 9, it can be concluded that for imperfect decision feedback the proposed EKF algorithm is not very sensitive to pilot spacing at low-to-medium SNR while at high SNR pilot spacing has a more significant impact on MIMO BER performance. This result is expected given that the proposed EKF inherently relies less on the observations and more on the state model at low SNRs. However, as the SNR increases the EKF algorithm updates the phase noise estimates by relying more and more on the observations.

Fig. 10 compares the BER performance of a 2×2 MIMO system for higher order modulations, i.e., QPSK or *16-quadrature amplitude modulation (16-QAM)*. The results in Fig. 10 show that even for denser constellations with imperfect decision feedback the proposed EKF is capable of tracking the phase noise over the frame and improving the overall system performance quite significantly. However, from the results in Fig. 10, one can deduce that compared to BPSK modulation, by employing QPSK or 16-QAM, the BER gap between a MIMO system based on perfect and imperfect decision feedback grows larger, since denser constellations are more susceptible to erroneous decoding in the presence of phase noise. This performance gap for denser constellations can be reduced by employing error-correcting codes in conjunction with the proposed EKF estimator.

Fig. 11 depicts the BER performance of 2×2 and 4×4 MIMO systems in the presence of imperfect and perfect channel and phase noise estimation. Fig. 11 shows that using the proposed channel and EKF estimators the BER performance of a 4×4 MIMO system is only 3 dB apart from the idealistic case of perfect channel and phase noise estimation at low-to-medium SNR. However, from the results in Fig. 11 it can be concluded that the performance of the proposed EKF estimator degrades as the dimensionality of the MIMO system increases. Therefore, for a 4×4 system, at high SNR, overall
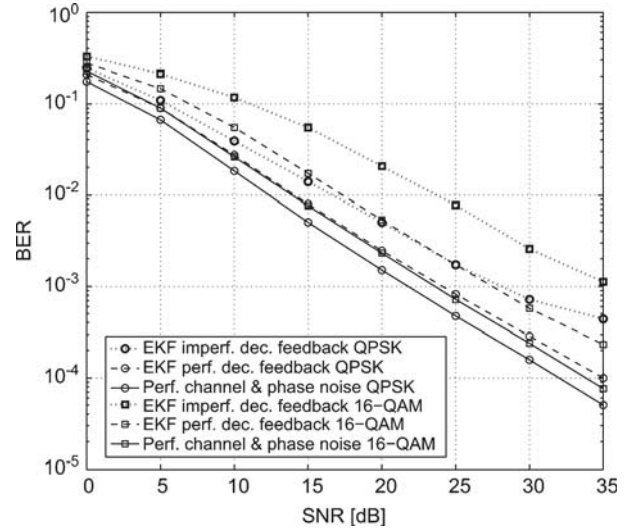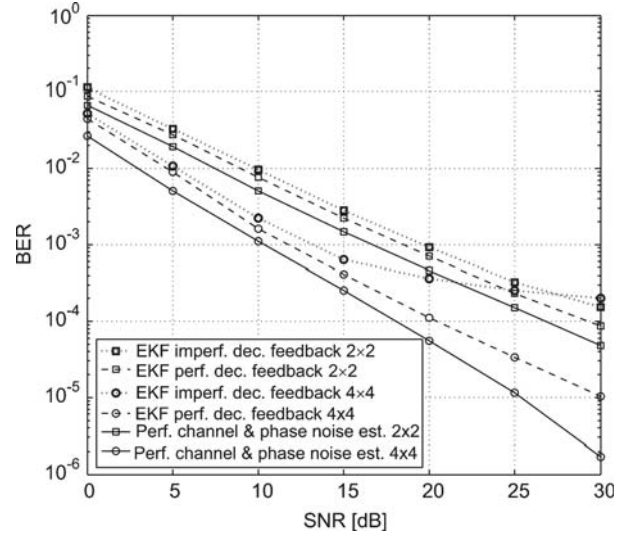

Fig. 11. Perfect and Imperfect decision feedback BER of a MIMO system for different N_t and N_r values, (BPSK modulation, $\sigma_\Delta^2 = 10^{-4} \, \mathrm{rad}^2$, and $L_p = 10$).

system performance is dominated by phase noise and compared to a 2×2 MIMO system, the performance gap between the cases of perfect and imperfect estimation widens. Note that for many practical applications, such as microwave backhauling, the number of transmit and receive antennas are not very large and the proposed EKF algorithm can be applied effectively to enable point-to-point communications over a wide range of SNR values.

Fig. 12 illustrates the perfect and imperfect decision feedback BER of a 2×2 MIMO system using the proposed EKF estimator for different Rician factors, $K = [-3, 2, 10]$ dB. The results in Fig. 12 demonstrate that the proposed EKF with imperfect decision feedback is capable of tracking the phase noise even for channels with a strong NLoS component, i.e., $K = -3$ dB. In addition, it is observed that the performance gap between the BER for perfect and imperfect decision feedback scenarios is reduced as the Rician factor K increases. This
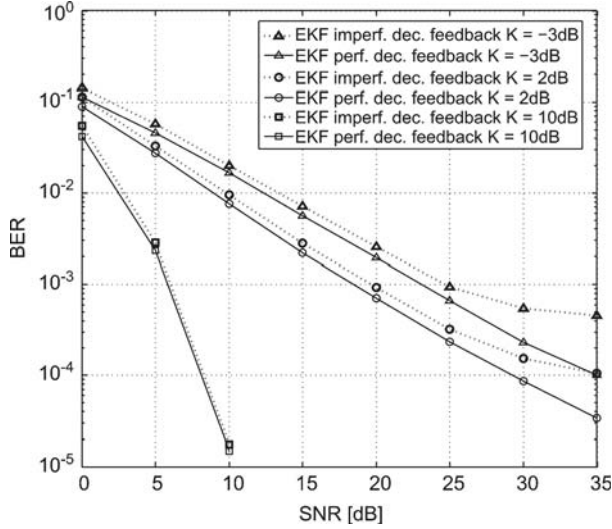
Fig. 12. Imperfect and perfect decision feedback BER of a 2×2 MIMO system for different Rician factors K, (BPSK, $L_p = 10$, and $\sigma_\Delta^2 = 10^{-4}$ rad^2).

Fig. 14. BER of a 2×2 MIMO system in the presence of time-varying and phase noise (BPSK and $\sigma_\Delta^2 = 10^{-4}$ rad^2).

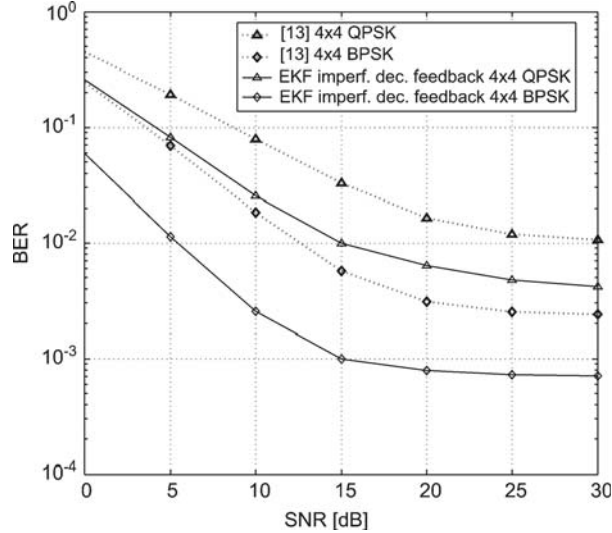

Fig. 13. BER performance of a 4×4 MIMO system using the proposed algorithm compared to that of [13] for QPSK and BPSK modulations ($L_p = 10$ and $\sigma_\Delta^2 = 10^{-3}$ rad^2).

result can be anticipated given that as K increases the MIMO channel quality also improves, resulting in a smaller bit error rate and in turn improving phase noise tracking in the case of imperfect decision feedback.

Fig. 13 compares the BER performance of a 4×4 MIMO system employing the proposed channel and phase noise estimation algorithm against that of [13]. As outlined in [13], $N_t = 4$ orthogonal pilot symbols need to be transmitted from each antenna. Thus, in order to ensure a fair comparison and maintain a 10% synchronization overhead, the results corresponding to [13] are generated by transmitting orthogonal pilot symbols before every stream of 40 data symbols. In addition, instead of assuming perfect channel knowledge as in [13], the MIMO channels are estimated using the transmitted pilot symbols for both algorithms. Fig. 13 shows that for the same synchronization overhead, the proposed algorithm noticeably outperforms the
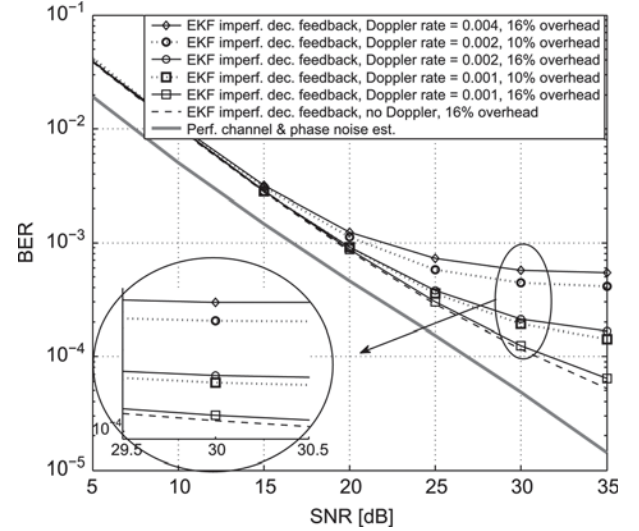
scheme in [13] for both QPSK and BPSK modulations. This result is expected since unlike the proposed EKF estimator, the approach in [13] does not provide any means of tracking the phase noise parameters using the received data symbols. Therefore, phase noise rotates the signal constellation and results in a significantly higher BER. Moreover, the results in Fig. 13 show that even though the performance of the proposed EKF based algorithm degrades as the number of antennas increases, it outperforms the approach in [13] for different modulations.

Fig. 14 depicts the BER performance of a MIMO system employing the proposed scheme in time-varying channels with Doppler rates[9] of $[0.001, 0.002, 0.004]$ and synchronization overheads of $[10\%, 16.6\%]$. The results in Fig. 14 show that even though the channels' time varying nature negatively affects system performance, for low Doppler rates and a synchronization overhead of 16.6%, a MIMO system's BER employing the proposed LS channel estimator and EKF phase noise tracking algorithm is close to that of quasi-static channels, i.e., no Doppler. However, as anticipated, as the Doppler rate increases a higher synchronization overhead is needed by the proposed scheme to maintain the same system performance. Therefore, Fig. 14 shows that by taking advantage of both training and received data symbols and balancing the tradeoff between synchronization overhead and performance, the proposed channel and phase noise tracking approach can track different channel conditions, e.g., high and low phase noise variances and Doppler rates, to achieve a specific system performance.

VI. CONCLUSION

In this paper, the estimation and effect of channel and phase noise in SDM MIMO systems is analyzed. After outlining the system model and deriving the CRLB for the multi-parameter estimation problem, a new data-aided LS algorithm is proposed that can jointly estimate the channel gains and phase noise. The MSE of the proposed LS estimator is shown to be close to the

---

[9]The time-varying channels are generated based on the Jakes' spectrum [63].

CRLB over a wide range of SNR values. In order to track the time-varying phase noise throughout a frame using the pilot and estimated data symbols, decision-directed WLS and EKF based estimators are proposed and it is shown that these estimators' MSEs are close to the CRLB. Next, the combination of the proposed LS and EKF estimators are applied to investigate the BER performance of an uncoded MIMO system in the presence of phase noise. The comparison is carried out for many different parameters such as phase noise rate, pilot-spacing, choice of modulation, and number of antennas. It is shown that the performance of a MIMO system using the proposed estimators is close to the idealistic setting of perfect channel and phase noise estimation. For example, at an SNR of 20 dB for a 2×2 MIMO system with imperfect decision feedback and a synchronization overhead of 10%, there is a performance gap of 3 dB between the two systems. We anticipate that this performance gap can be further reduced by performing soft-decision feedback and phase tracking using all the symbols within a frame, i.e., off-line processing. In addition, the results in Section V show that by selecting an appropriate synchronization overhead, the proposed estimators can track MIMO channels and phase noise processes in the presence of various Doppler rates. These results demonstrate that the proposed channel and phase noise tracking schemes can be used to enable application of MIMO systems to new frontiers such as point-to-point microwave backhaul and satellite communication links. Note that even though the proposed scheme cannot be directly applied to frequency selective channels, the principles and methodologies proposed here can be used to develop new channel and multiple phase noise estimation algorithms for such channels. For example, it is well-known that application of OFDM and *orthogonal frequency division multiple access (OFDMA)* can significantly improve the performance of communication systems in frequency selective channels. However, similar to single carrier systems, OFDM and OFDMA systems may be affected by multiple multiplicative phase noise processes [15], i.e., CPEs. Therefore, the algorithms proposed here can be modified and applied to estimate multiple phase noise parameters in such systems and improve their performance. However, addressing this specific problem is beyond the scope of this paper and can be the subject of future work.

APPENDIX
DERIVATION OF FIM

Here, the FIM for the data-aided estimation of $\boldsymbol{\lambda}$ is derived. The received signal vector at the kth receive antenna, $\mathbf{y}_k \triangleq [y_k(n - L_t + 1), \ldots, y_k(n)]^T$ is given by

$$\mathbf{y}_k = \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell$$

$$- j \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell \odot \boldsymbol{\psi}_{k,\ell} + \mathbf{w}_k \quad \text{(A.1)}$$

where

- $\boldsymbol{\psi}_{k,\ell} \triangleq \left[\sum_{m=n-L_t+2}^{n} \Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m), \sum_{m=n-L_t+3}^{n} \right.$
$\left. \times \Delta_k^{[r]}(m) + \Delta_\ell^{[t]}(m), \ldots, 0 \right]^T$;

- $\mathbf{s}_\ell \triangleq [s_\ell(n - L_t + 1), \ldots, s_\ell(n)]^T$;
- $\mathbf{w}_k \triangleq [w_k(n - L_t + 1), \ldots, w_k(n)]^T$.

Based on the assumptions in Section II, the vector of received signals at all receive antennas, $\mathbf{y} \triangleq \left[\mathbf{y}_1^T, \ldots, \mathbf{y}_{N_r}^T \right]^T$, is distributed as $\mathbf{y} \sim \mathcal{CN}\left(\boldsymbol{\mu}_{\mathbf{y}}, \boldsymbol{\Sigma}_{\mathbf{y}} \right)$, where $\boldsymbol{\mu}_{\mathbf{y}} \triangleq \left[\boldsymbol{\mu}_{\mathbf{y}_1}^T, \ldots, \boldsymbol{\mu}_{\mathbf{y}_{N_r}}^T \right]^T$ with $\boldsymbol{\mu}_{\mathbf{y}_k} = \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell$, and the $L_t \times L_t$ submatrices $\boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}$ of the $N_r L_t \times N_r L_t$ covariance matrix, $\boldsymbol{\Sigma}_{\mathbf{y}}$, for $k, \bar{k} = 1, \ldots, N_r$, can be determined as

$$\boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}} = \mathbb{E}\left[\left(-j \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} \mathbf{s}_\ell \odot \boldsymbol{\psi}_{k,\ell} + \mathbf{w}_k \right) \right.$$
$$\left. \times \left(-j \sum_{\bar{\ell}=1}^{N_t} \alpha_{\bar{k},\bar{\ell}} e^{j\beta_{\bar{k},\bar{\ell}}(n)} \mathbf{s}_{\bar{\ell}} \odot \boldsymbol{\psi}_{\bar{k},\bar{\ell}} + \mathbf{w}_{\bar{k}} \right)^H \right] \quad \text{(A.2a)}$$

$$= \begin{cases} \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{k,\ell} \alpha_{\bar{k},\bar{\ell}} e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\bar{\ell}}(n)} \\ \times \mathbb{E}\left[\left(\mathbf{s}_\ell \odot \boldsymbol{\psi}_{k,\ell} \right) \left(\mathbf{s}_{\bar{\ell}} \odot \boldsymbol{\psi}_{\bar{k},\bar{\ell}} \right)^H \right] \\ + \sigma_w^2 \mathbf{I}_{L_t \times L_t}, \quad\quad\quad\quad\quad k = \bar{k} \\ \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{k,\ell} \alpha_{\bar{k},\bar{\ell}} e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\bar{\ell}}(n)} \\ \times \mathbb{E}\left[\left(\mathbf{s}_\ell \odot \boldsymbol{\psi}_{k,\ell} \right) \left(\mathbf{s}_{\bar{\ell}} \odot \boldsymbol{\psi}_{\bar{k},\bar{\ell}} \right)^H \right], \quad k \neq \bar{k} \end{cases} \quad \text{(A.2b)}$$

$$= \begin{cases} \sum_{\ell=1}^{N_t} \alpha_{k,\ell}^2 \mathbf{s}_\ell \mathbf{s}_\ell^H \\ \odot (\sigma_{\Delta_k^{[r]}}^2 \boldsymbol{\Psi} + \sigma_{\Delta_\ell^{[t]}}^2 \boldsymbol{\Psi}) + \sum_{\ell=1}^{N_t} \sum_{\substack{\bar{\ell}=1, \\ \bar{\ell} \neq \ell}}^{N_t} \alpha_{k,\ell} \alpha_{\bar{k},\bar{\ell}} \\ \times e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\bar{\ell}}(n)} \mathbf{s}_\ell \mathbf{s}_\ell^H \\ \odot (\sigma_{\Delta_k^{[r]}}^2 \boldsymbol{\Psi}) + \sigma_w^2 \mathbf{I}_{L_t \times L_t}, \quad\quad k = \bar{k} \\ \sum_{\ell=1}^{N_t} \alpha_{k,\ell} \alpha_{\bar{k},\ell} e^{j\beta_{k,\ell}(n)} e^{-j\beta_{\bar{k},\ell}(n)} \\ \times \mathbf{s}_\ell \mathbf{s}_\ell^H \odot (\sigma_{\Delta_\ell^{[t]}}^2 \boldsymbol{\Psi}), \quad\quad\quad\quad k \neq \bar{k}. \end{cases} \quad \text{(A.2c)}$$

Note that (A.2b) follows from (A.2a) due to the assumption that the AWGN and phase noise innovations $\Delta_\ell^{[t]}, \Delta_k^{[r]}, \forall k, \ell$, are mutually independent. Moreover, (A.2c) follows from the assumption of mutual independence between the phase noise innovations corresponding to difference symbols[10]. Given that the observation sequence, $\mathbf{y}$, has a complex Gaussian distribution, the ith row and $\bar{i}$th column entry of the submatrix $\mathbf{FIM}_{k,\bar{k}}$ for $\ell, \bar{\ell} = 1, \ldots, 2N_t$ and $k, \bar{k} = 1, \ldots, N_r$ is given by [61]

$$\left[\mathbf{FIM}_{k,\bar{k}}\right]_{\ell,\bar{\ell}} = 2\Re \left\{ \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \lambda_{k,\ell}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \lambda_{\bar{k},\bar{\ell}}} \right\}$$
$$+ \text{Tr} \left(\boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \lambda_{k,\ell}} \boldsymbol{\Sigma}_{\mathbf{y}}^{-1} \frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \lambda_{\bar{k},\bar{\ell}}} \right), \quad \text{(A.3)}$$

where $\lambda_{k,\ell}$ denotes the ℓth element of the vector of parameters of interests corresponding to the kth receive antenna, $\boldsymbol{\lambda}_k$ in

[10]$\Delta_k^{[r]}(n)$ and $\Delta_k^{[r]}(m)$ are assumed to be mutually independent for $m \neq n$.

$$
\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}}{\partial \alpha_{m,\ell}} = \begin{cases} 2\alpha_{m,\ell}\mathbf{s}_\ell\mathbf{s}_\ell^H \odot (\sigma^2_{\Delta_m^{[r]}}\boldsymbol{\Psi} + \sigma^2_{\Delta_\ell^{[t]}}\boldsymbol{\Psi}) + \sum_{p=1,p\neq\ell}^{N_t}\left(\alpha_{m,p}e^{j\beta_{m,\ell}(n)}e^{-j\beta_{m,p}(n)}\right. \\ \quad \times\mathbf{s}_\ell\mathbf{s}_p^H \odot \sigma^2_{\Delta_m^{[r]}}\boldsymbol{\Psi} + \alpha_{m,p}e^{j\beta_{m,p}(n)}e^{-j\beta_{m,\ell}(n)}\mathbf{s}_p\mathbf{s}_\ell^H \odot \left. \sigma^2_{\Delta_m^{[r]}}\boldsymbol{\Psi}\right), & m = k = \bar{k} \\ \alpha_{\bar{k},\ell}e^{j\beta_{m,\ell}(n)}e^{-j\beta_{\bar{k},\ell}(n)}\mathbf{s}_\ell\mathbf{s}_\ell^H \odot \sigma^2_{\Delta_\ell^{[t]}}\boldsymbol{\Psi}, & m = k,\ \bar{k} \neq k \\ \alpha_{k,\ell}e^{j\beta_{k,\ell}(n)}e^{-j\beta_{m,\ell}(n)}\mathbf{s}_\ell\mathbf{s}_\ell^H \odot \sigma^2_{\Delta_\ell^{[t]}}\boldsymbol{\Psi}, & m = \bar{k},\ k \neq \bar{k} \\ 0, & m \neq k,\ m \neq \bar{k} \end{cases} \tag{A.5}
$$

$$
\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}}{\partial \beta_{m,\ell}(n)} = \begin{cases} \sum_{p=1,p\neq\ell}^{N_t}\left(\alpha_{m,\ell}\alpha_{m,p}je^{j\beta_{m,\ell}(n)}e^{-j\beta_{m,p}(n)}\mathbf{s}_\ell\mathbf{s}_p^H \odot \sigma^2_{\Delta_m^{[r]}}\boldsymbol{\Psi}\right. \\ \quad +\alpha_{m,p}\alpha_{m,\ell}(-j)e^{j\beta_{m,p}(n)}e^{-j\beta_{m,\ell}(n)}\mathbf{s}_p\mathbf{s}_\ell^H \odot \left.\sigma^2_{\Delta_m^{[r]}}\boldsymbol{\Psi}\right), & m = k = \bar{k} \\ \alpha_{m,\ell}\alpha_{\bar{k},\ell}je^{j\beta_{m,\ell}(n)}e^{-j\beta_{\bar{k},\ell}(n)}\mathbf{s}_\ell\mathbf{s}_\ell^H \odot \sigma^2_{\Delta_\ell^{[t]}}\boldsymbol{\Psi}, & m = k,\ \bar{k} \neq k \\ \alpha_{m,\ell}\alpha_{k,\ell}(-j)e^{j\beta_{k,\ell}(n)}e^{-j\beta_{m,\ell}(n)}\mathbf{s}_\ell\mathbf{s}_\ell^H \odot \sigma^2_{\Delta_\ell^{[t]}}\boldsymbol{\Psi}, & m = \bar{k},\ k \neq \bar{k} \\ 0, & m \neq k,\ m \neq \bar{k}. \end{cases} \tag{A.6}
$$

(A.1). In order to evaluate the FIM, the derivatives $\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \alpha_{k,\ell}}$ and $\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \beta_{k,\ell}(n)}$ need to be determined, where

$$
\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \alpha_{k,\ell}} = \Big[\underbrace{0,\ldots,0}_{L_t(k-1)}, e^{j\beta_{k,\ell}(n)}\mathbf{s}_\ell^T, \underbrace{0,\ldots,0}_{L_t(N_r-k)}\Big]^T,
$$

$$
\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \beta_{k,\ell}(n)} = \Big[\underbrace{0,\ldots,0}_{L_t(k-1)}, \alpha_{k,\ell}je^{j\beta_{k,\ell}(n)}\mathbf{s}_\ell^T, \underbrace{0,\ldots,0}_{L_t(N_r-k)}\Big]^T. \tag{A.4}
$$

In addition, the derivatives of the sub-matrices $\boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}$ in (A.2), $\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}}{\partial \alpha_{m,\ell}}$ and $\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}_{k,\bar{k}}}}{\partial \beta_{m,\ell}(n)}$, for $m = 1,\ldots,N_r$, can be evaluated as shown in (A.5) and (A.5) at the top of the page. After substituting the derivatives in (A.4)–(A.6) into (A.3) and carrying out straightforward algebraic manipulations, the elements of $\mathbf{FIM}_{k,\bar{k}}$ can be obtained as

$$
\left[\mathbf{FIM}_{k,\bar{k}}\right]_{\ell,\bar{\ell}} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \alpha_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}}\right\}
$$
$$
+ \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}}\right\}, \tag{A.7a}
$$

$$
\left[\mathbf{FIM}_{k,\bar{k}}\right]_{\ell,\bar{\ell}+N_t} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \alpha_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}}\right\}
$$
$$
+ \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}}\right\}, \tag{A.7b}
$$

$$
\left[\mathbf{FIM}_{k,\bar{k}}\right]_{\ell+N_t,\bar{\ell}} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \beta_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}}\right\}
$$
$$
+ \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \alpha_{\bar{k},\bar{\ell}}}\right\}, \tag{A.7c}
$$

$$
\left[\mathbf{FIM}_{k,\bar{k}}\right]_{\ell+N_t,\bar{\ell}+N_t} = 2\Re\left\{\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}^H}{\partial \beta_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\mu}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}}\right\}
$$

$$
+ \mathrm{Tr}\left\{\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{k,\ell}}\boldsymbol{\Sigma}_{\mathbf{y}}^{-1}\frac{\partial \boldsymbol{\Sigma}_{\mathbf{y}}}{\partial \beta_{\bar{k},\bar{\ell}}}\right\} \tag{A.7d}
$$

Using (A.7a)–(A.7d), $\mathbf{FIM}_{k,\bar{k}}$, for k, $\bar{k} = 1,\ldots,N_r$, can be obtained as shown in (14) in Section III-A.

REFERENCES

[1] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Commun.*, vol. 10, no. 6, pp. 585–595, Nov.-Dec. 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, Mar. 1998.

[3] H. Meyr, M. Moeneclaey, and S. A. Fechtel, "Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing," in *Wiley-InterScience*. New York: Wiley, 1997.

[4] M. Dohler, R. W. Heath, A. Lozano, C. B. Papadias, and R. A. Valenzuela, "Is the PHY layer dead?," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 159–165, Apr. 2011.

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

[6] G. Klimovitch, "A nonlinear theory of near-carrier phase noise in free running oscillators," in *Proc. IEEE Int. Conf. Circuits Syst.*, Mar. 2000, pp. 1–6.

[7] L. Tomba, "On the effect of Wiener phase noise in OFDM systems," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 580–583, May 1998.

[8] F. Herzel, "An analytical model for the power spectral density of a voltage controlled oscillator and its analogy to the laser linewidth theory," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 45, no. 9, pp. 904–908, Sep. 1998.

[9] G. Niu, "Noise in SiGe HBT RF technology: Physics, modeling, and circuit implications," *Proc. IEEE*, vol. 93, no. 9, pp. 1583–1597, Sep. 2005.

[10] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst.*, vol. 47, no. 5, pp. 655–674, May 2000.

[11] J. A. McNeill, "Jitter in ring oscillators," *IEEE J. Solid-State Circuits*, vol. 32, no. 6, pp. 870–879, Jun. 1997.

[12] A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," *IEEE J. Solid-State Circuits*, vol. 34, no. 6, pp. 790–804, Jun. 1999.

[13] N. Hadaschik, M. Dorpinghaus, A. Senst, O. Harmjanz, and U. Kaufer, "Improving MIMO phase noise estimation by exploiting spatial correlations," in *Proc. IEEE Acoust., Speech, Signal Process. (ICASSP)*, Philadelphia, PA, Mar. 2005, vol. 3.

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

[15] Y. Wang and D. Falconer, "Phase noise estimation and suppression for single carrier SDMA," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Sydney, Australia, Jul. 2010.

[16] S. A. Merritt, "The iterated extended Kalman phase detector," *IEEE Trans. Commun.*, vol. 37, no. 5, pp. 522–526, May 1989.

[17] L. Zhao and W. Namgoong, "A novel phase-noise compensation scheme for communication receivers," *IEEE Trans. Commun.*, vol. 54, no. 3, pp. 532–542, Mar. 2006.

[18] V. Simon, A. Senst, M. Speth, and H. Meyr, "Phase noise estimation via adapted interpolation," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, San Antonio, TX, Dec. 2001, vol. 6, pp. 3297–3301.

[19] J. Bhatti and M. Moeneclaey, "Feedforward data-aided phase noise estimation from a DCT basis expansion," *EURASIP J. Wireless Commun. Netw., Special Issue on Synch. Wireless Commun.*, Jan. 2009.

[20] J. Dauwels and H.-A. Loeliger, "Joint decoding and phase estimation: An exercise in factor graphs," in *Proc. IEEE Inter. Symp. Inf. Theory*, Yokohama, Japan, Jul. 2003, pp. 231–231.

[21] P. Amblard, J. Brossier, and E. Moisan, "Phase tracking: What do we gain from optimality? particle filtering versus phase-locked loops," *Elsevier J. Signal Process.*, vol. 83, no. 1, pp. 151–167, 2003.

[22] S. Bay, C. Herzet, J.-M. Brossier, J.-P. Barbot, and B. Geller, "Analytic and asymptotic analysis of Bayesian Cramér-Rao bound for dynamical phase offset estimation," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 61–70, Jan. 2008.

[23] N. Noels, H. Steendam, M. Moeneclaey, and H. Bruneel, "Carrier phase and frequency estimation for pilot-symbol assisted transmission: Bounds and algorithms," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4578–4587, Dec. 2005.

[24] S. Godtmann, N. Hadaschik, A. Pollok, G. Ascheid, and H. Meyr, "Iterative code-aided phase noise synchronization based on the LMMSE criterion," in *Proc. IEEE Workshop on Signal Process. Adv. Wireless Commun.*, Helsinki, Finland, Jun. 2007, pp. 1–5.

[25] T. S. Shehata and M. El-Tanany, "Joint iterative detection and phase noise estimation algorithms using Kalman filtering," in *Proc. Canad. Workshop on Inf. Theory*, Ottawa, Canada, May 2009.

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

[27] J. Bhatti and M. Moeneclaey, "Performance analysis of iterative decision-directed phase noise estimation," in *Proc. Future Netw. Mobile Summit*, Florence, Italy, Jun. 2010, pp. 1–8.

[28] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP J. Wireless Commun. Netw., Special Issue on Adv. Signal Process. Algorithms for Wireless Commun.*, vol. 2005, no. 2, pp. 117–129, Apr. 2005.

[29] M. Nissilä and S. Pasupathy, "Adaptive iterative detectors for phase-uncertain channels via variational bounding," *IEEE Trans. Commun.*, vol. 57, no. 3, pp. 716–715, Mar. 2009.

[30] M. R. Khanzadi *et al.*, "On models, bounds, and estimation algorithms for time-varying phase noise," in *Proc. Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, Dec. 2011.

[31] T. Höhne and V. Ranki, "Phase noise in beamforming," *IEEE Trans. Wireless Commun.*, vol. 9, no. 12, pp. 1696–1705, Dec. 2010.

[32] A. Taparugssanagorn and J. Ylitalo, "Characteristics of short-term phase noise of MIMO channel sounding and its effect on capacity estimation," *IEEE Trans. Instrum. Meas.*, vol. 58, no. 1, pp. 196–201, Jan. 2009.

[33] D. Baum and H. Bolcskei, "Impact of phase noise on MIMO channel measurement accuracy," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Los Angeles, CA, Sep. 2004, vol. 3, pp. 1614–1618.

[34] T. C. W. Schenk, X.-J. Tao, P. F. M. Smulders, and E. R. Fledderus, "Influence and suppression of phase noise in multi-antenna OFDM," in *Proc. IEEE Veh. Technol. Conf. (VTC)*, Los Angeles, CA, Sep. 2004, vol. 2, pp. 1443–1447.

[35] 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.

[36] A. G. Armada, "Understanding the effects of phase noise in orthogonal frequency division multiplexing (OFDM)," *IEEE Trans. Broadcast.*, vol. 47, no. 2, pp. 153–159, Jun. 2001.

[37] P. Liu, S. Wu, and Y. Bar-Ness, "A phase noise mitigation scheme for MIMO WLANS with spatially correlated and imperfectly estimated channels," *IEEE Commun. Lett.*, vol. 10, no. 3, pp. 141–143, Mar. 2006.

[38] T. Schenk, *RF Imperfections in High-Rate Wireless Systems*. New York: Springer, 2008.

[39] A. G. Armada and M. Calvo, "Phase noise and sub-carrier spacing effects on the performance of an OFDM communication system," *IEEE Commun. Lett.*, vol. 2, no. 1, pp. 11–13, Jan. 1998.

[40] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Performance analysis of OFDM with Wiener phase noise and frequency selective fading channels," *IEEE Trans. Commun.*, vol. 59, no. 5, pp. 1321–1331, May 2011.

[41] X. Zhang and H.-G. Ryu, "Joint estimation and suppression of phase noise and carrier frequency offset in multiple-input multiple-output single carrier frequency division multiple access with single-carrier space frequency block coding," *IE Trans. Commun.*, vol. 4, no. 16, pp. 1998–2007, 2010.

[42] D. D. Lin, R. Pacheco, T. J. Lim, and D. Hatzinakos, "Joint estimation of channel response, frequency offset, and phase noise in OFDM," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3442–3554, Sep. 2006.

[43] D. D. Lin and T. J. Lim, "The variational inference approach to joint data detection and phase noise estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1862–1874, May 2007.

[44] T. Pollet, M. V. Blade, and M. Moeneclaey, "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 191–193, Feb./Mar./Apr. 1995.

[45] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[46] G. E. Corazza, *Digital Satellite Communications*. New York: Springer Science plus Business Media, LLC, 2007.

[47] J. Hansryd and J. Edstam, Microwave Capacity Evolution: New Technologies and Additional Frequencies Ericsson AB Publications, 2011 [Online]. Available: http://www.ericsson.com/news/110621_microwave_capacity_evolution_c

[48] T. Manning, *Microwave Radio Transmission Design Guide*, 2nd ed. Boston, MA: Artech House, 2009.

[49] J. Wells, *Multi-Gigabit Microwave and Millimeter-Wave Wireless Communications*, 1st ed. Boston, MA: Artech House, 2010.

[50] K.-C. Huang and D. J. Edwards, *Millimetre Wave Antennas for Gigabit Wireless Communications: A Practical Guide to Design and Analysis in a System Context*. New York: Wiley, 2008.

[51] H. Mehrpouyan *et al.*, "Time-varying phase noise and channel estimation in MIMO systems," in *Proc. IEEE Int. Workshop on Signal Process. Adv. Wireless Commun. (SPWAC)*, Jun. 2012.

[52] N. A. D'Andrea, U. Mengali, and R. Reggiannini, "Comparison of carrier recovery methods for narrow-band polyphase shift keyed signals," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Hollywood, FL, Dec. 1988, vol. 3, pp. 1474–1478.

[53] M. Moeneclaey and G. de Jonghe, "Ml-oriented NDA carrier synchronization for general rotationally symmetric signal constellations," *IEEE Trans. Commun.*, vol. 42, no. 8, pp. 2531–2533, Aug. 1994.

[54] B. G. Quinn, "Estimating frequency by interpolation using Fourier coefficients," *IEEE Trans. Signal Process.*, vol. 42, no. 5, pp. 1264–1268, May 1994.

[55] S. M. Kay, "Fundamentals of Statistical Signal Processing, Estimation Theory," in *Signal Processing Series*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[56] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York: Wiley, 2001.

[57] Y. Nasser, M. des Noes, L. Ros, and G. Jourdain, "On the system level prediction of joint time frequency spreading systems with carrier phase noise," *IEEE Trans. Commun.*, vol. 58, no. 3, pp. 839–850, Mar. 2010.

[58] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[59] H. Mehrpouyan and S. D. Blostein, "Bounds and algorithms for multiple frequency offset estimation in cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1300–1311, Apr. 2011.

[60] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.

[61] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[62] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Boston, MA: Kluwer Academic, 2001.

[63] H. Mehrpouyan and S. D. Blostein, "ARMA synthesis of fading channels," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 2846–2850, Jul. 2007.

Hani Mehrpouyan (S'05–M'10) received the B.Sc. honors degree in computer engineering from Simon Fraser University, Burnaby, Canada, in 2004. In 2005, he joined the Department of ECE, Queen's University, Kingston, Canada, as a Ph.D. candidate, where he received the Ph.D. degree in electrical engineering in 2010.

From September 2010 to March of 2012, he was a Postdoctoral Researcher with the Department of Signal and Systems, Chalmers University of Technology, where he led the MIMO aspects of the microwave backhauling for next generation wireless networks project. He next joined the University of Luxembourg as a Research Associate from April 2012 to August of 2012, where he was responsible for new interference cancelation and synchronization schemes for next generation satellite communication links. Since August 2012, he has been an Assistant Professor with the Department of

Computer and Electrical Engineering and Computer Science, California State University, Bakersfield. His current research interests lie in the area of applied signal processing and physical layer of wireless communication systems, including synchronization, channel estimation, interference cancelation, and performance optimization.

Dr. Mehrpouyan has received more than 10 scholarships and awards, e.g., IEEE Globecom Early Bird Student Award, NSERC-IRDF, NSERC PGS-D, NSERC CGS-M Alexander Graham Bell, B.C. Wireless Innovation, and more. He has more than 20 publications in prestigious IEEE Journals and Conferences. He has served as a TPC member for IEEE Globecom 2012. He has also been involved with industry leaders such as Ericsson AB, Research in Motion (RIM), Alcatel, etc. For more information, refer to www.mehrpouyan.info.

Ali A. Nasir (S'11) received the B.Sc. (1st class hons.) degree in electrical engineering from the University of Engineering and Technology (UET), Lahore, Pakistan, in 2007.

He is currently pursuing the Ph.D. degree at the Research School of Engineering, the Australian National University (ANU), Canberra. Prior to joining ANU, he worked as a Lecturer in UET for nine months and then as a Design Engineer in the Centre for Advanced Research in Engineering (CARE), Islamabad, Pakistan, for a year. He was a Research Visitor in Chalmers University of Technology, Gothenburg, Sweden, for six months from April to October 2011. His research interests are in the area of synchronization and channel estimation in cooperative communication, MIMO, and OFDM systems.

Mr. Nasir was awarded a University Gold Medal for outstanding performance during the final year of his undergraduate studies. He is a recipient of an ANU International Ph.D. scholarship for the duration of his Ph.D. He was also awarded an ANU Vice Chancellor's Higher Degree Research (HDR) travel grant in 2011.

Steven D. Blostein (S'83–M'88–SM'96) received the B.S. degree in electrical engineering from Cornell University, Ithaca, NY, in 1983, and the M.S. and Ph.D. degrees in electrical and computer engineering, University of Illinois, Urbana-Champaign, in 1985 and 1988, respectively.

He has been on the faculty of the Department of Electrical and Computer Engineering, Queen's University, since 1988 and currently holds the position of Professor. From 2004 to 2009, he was Department Head. From 1999 to 2003, he was the leader of the Multi-Rate Wireless Data Access Major Project sponsored by the Canadian Institute for Telecommunications Research. He has also been a consultant to industry and government in the areas of image compression, target tracking, radar imaging, and wireless communications. He spent sabbatical leaves at Lockheed Martin Electronic Systems and at Communications Research Centre, Ottawa, Canada. His current interests lie in the application of signal processing to problems in wireless communications systems, including synchronization, network MIMO, and physical layer optimization for multimedia transmission. He has been a member of the Samsung 4G Wireless Forum as well as an invited distinguished speaker.

Dr. Blostein served as Chair of IEEE Kingston Section (1994), Chair of the Biennial Symposium on Communications (2000, 2006, 2008), Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING (1996–2000), and Publications Chair for IEEE ICASSP 2004. For a number of years, he has been serving on Technical Program Committees for the IEEE Communications Society conferences that include ICC, Globecom, and WCNC. He has been serving as an editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS since 2007. He is a registered professional engineer in Ontario.

Thomas Eriksson received the Ph.D. degree in information theory in 1996 from Chalmers University of Technology, Gothenburg, Sweden.

From 1990 to 1996, he was with Chalmers. From 1997 to 1998, he was with AT&T Labs—Research, Murray Hill, NJ, and in 1998 and 1999, he was with Ericsson Radio Systems AB, Kista, Sweden. Since 1999, he has been with Chalmers University, where he is currently a professor in communication systems. Further, he was a guest professor at Yonsei University, South Korea, during 2003–2004. His research interests include communication, data compression, and modeling and compensation of nonideal hardware components (e.g., amplifiers and oscillators in communication transmitters and receivers).

George K. Karagiannidis (SM'03) was born in Pithagorion, Samos Island, Greece. He received the University Diploma (5 years) and Ph.D. degree, both in electrical and computer engineering, from the University of Patras, Greece, in 1987 and 1999, respectively.

From 2000 to 2004, he was a Senior Researcher at the Institute for Space Applications and Remote Sensing, National Observatory of Athens, Greece. In June 2004, he joined the faculty of Aristotle University of Thessaloniki, Greece, where he is currently Associate Professor (four-level academic rank system) of Digital Communications Systems at the Electrical and Computer Engineering Department and Head of the Telecommunications Systems and Networks Lab. His research interests are in the broad area of digital communications systems with emphasis on communications theory, energy efficient MIMO and cooperative communications, cognitive radio, wireless security, and optical wireless communications. He is the author or coauthor of more than 160 technical papers published in scientific journals and presented at international conferences. He is also the author of the Greek edition of a book on telecommunications systems and coauthor of *Advanced Wireless Communications Systems* (Cambridge, U.K.: Cambridge Publications, 2012).

Dr. Karagiannidis is a corecipient of the Best Paper Award of the Wireless Communications Symposium (WCS) in the IEEE International Conference on Communications (ICC'07), Glasgow, U.K., June 2007. He has been a member of Technical Program Committees for several IEEE conferences such as ICC, GLOBECOM, VTC, etc. In the past, he was Editor for Fading Channels and Diversity of the IEEE TRANSACTIONS ON COMMUNICATIONS, Senior Editor of the IEEE COMMUNICATIONS LETTERS, and Editor of the EURASIP *Journal of Wireless Communications & Networks*. He was Lead Guest Editor of the special issue on "Optical Wireless Communications" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and currently serves as a Guest Editor of the special issue on "Large-scale multiple antenna wireless systems" (to be published in February 2013). Since January 2012, he has been the Editor-in Chief of the IEEE COMMUNICATIONS LETTERS.

Tommy Svensson (SM'09) received the M.S. degree in engineering physics 1994 and the Ph.D. degree in information theory 2003, both from Chalmers University of Technology, Sweden.

He has contributed to the development of the 3GPP LTE standard and its evolution through the European collaboration projects EU FP6 WINNER, WINNER II, CELTIC WINNER+, and ARTIST4G. He is currently an Associate Professor in Communication Systems at Chalmers, where he is leading the research on air interface and microwave backhauling technologies for future wireless systems. He has also been working at Ericsson AB with core networks, radio access networks, and microwave transmission products. His main research interests are in design and analysis of physical layer algorithms (channel coding, modulation, and detection with awareness of hardware imperfections), multiple access schemes, resource allocation, cooperative systems, and moving relays, but he is also experienced in higher layer design of wireless communication systems.

Dr. Svensson is chairman of the Swedish IEEE VT/COM/IT chapter.