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—Multi-input multi-output (MIMO), Wiener phase noise, channel estimation, Cramér-Rao Lower Bound (CRLB), weighted least squares (WLS), and extended Kalman filter (EKF).

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 the 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 systems1, 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 multi-user 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

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

2For 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].
in [16], [20], [25], [29] are focused on signal detection and do not analyze or investigate the estimation performance of the 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 transmit 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 in 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 non-linear 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
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 MIMO systems with imperfect knowledge 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 to 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. Finally, Section VI concludes the paper and summarizes its key findings.

D. Notation

Superscripts $(\cdot)^*$, $(\cdot)^H$, and $(\cdot)^T$ denote the conjugate, the conjugate transpose, and the transpose operators, respectively. Bold face small letters, e.g., $x$, are used for vectors, bold face capital alphabets, e.g., $X$, are used for matrices, and $[X]_{x,y}$ represents the entry in row $x$ and column $y$ of $X$. $\textbf{1}_{X \times X}$, $\textbf{0}_{X \times X}$, and $\textbf{1}_{X \times X}$ 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 the complex variable $x$, $| x |$ denotes the element-wise absolute value of a vector $x$, $\text{diag}(x)$ is used to denote a diagonal matrix, where the diagonal elements are given by vector $x$. $\text{diag}(X)$ is used to denote the diagonal elements of matrix $X$. $\mathbb{E}[\cdot]$ 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, $\mathcal{N}(\mu, \sigma^2)$ and $\mathcal{CN}(\mu, \sigma^2)$ denote real and complex Gaussian distributions with mean $\mu$ and variance $\sigma^2$, 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 the literature [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.

---

\(^3\)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].
**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 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 $k$th antenna of the MIMO receiver is given by

$$
y_k(n) = \sum_{\ell=1}^{N_t} e^{j\theta_k^{\ell}(n)} h_{k,\ell} e^{j\theta_k^{\ell}(n)} s_{\ell}(n) + w_k(n)$$

where

- $s_{\ell}(n)$ is the $n$th, for $n = 1, \cdots, L_f$, $M$-ary modulated transmitted symbol that corresponds to the $\ell$th transmit antenna and consists of both pilots and data symbols,
- $h_{k,\ell}$ is the quasi-static *unknown* channel gain from the $\ell$th transmit to the $k$th 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., $h_{k,\ell} \sim \mathcal{C}\mathcal{N}(\mu_{h_{k,\ell}}, \sigma_{h_{k,\ell}}^2)$ from frame to frame,
- $\alpha_{k,\ell}$ is the channel gain from the $\ell$th transmit to the $k$th receive antenna,
- $\theta_k^{\ell}(n)$ and $\theta_k^{\ell}(n)$ correspond to the $n$th sample of the phase noise at the $\ell$th transmit and $k$th receive antenna, respectively,
- $\beta_{k,\ell}(n) \triangleq \theta_k^{\ell}(n) + \theta_k^{\ell}(n) + \angle h_{k,\ell}$ denotes the overall phase shift from the oscillator and channel corresponding to the $\ell$th transmit and $k$th receive antenna, and
- $w_k(n)$ is the zero-mean complex *additive white Gaussian noise (AWGN)* at the $k$th receive antenna, i.e., $w_k(n) \sim \mathcal{C}\mathcal{N}(0, \sigma_w^2(n))$. Note that the AWGN variance, $\sigma_w^2(n)$, 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, $\theta_k^{\ell} + \alpha_{k,\ell}$, for $\ell = 1, \cdots, N_t$ and $k = 1, \cdots, N_r$, are given by [5]–[10]

$$\theta_k^{\ell}(n) = \theta_k^{\ell}(n-1) + \Delta_k^{\ell}(n), \quad \text{and}$$

$$\theta_k^{\ell}(n) = \theta_k^{\ell}(n-1) + \Delta_k^{\ell}(n), \quad \text{(2)}$$

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

$$\sigma_{\Delta_k^{\ell}}^2 = 2\pi c_k^{\ell} T_s, \quad \text{and} \quad \sigma_{\Delta_k^{\ell}}^2 = 2\pi c_k^{\ell} T_s, \quad \text{(3)}$$

where the constants $c_k^{\ell}$ and $c_k^{\ell}$ denote the one-sided 3-dB bandwidth of the Lorentzian spectrum of the oscillators at the $\ell$th and $k$th 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 $\sigma_\Delta^2[\tau]$ and $\sigma_\Delta^2[\tau]$ are known at the receiver, given that they are dependent on the oscillator properties.

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

$$y(n) = \Theta^r(n)H\Theta^t(n)s(n) + w(n) = \left( A \odot e^{jB(n)} \right) s(n) + w(n),$$

(4)

where

- $y(n) \triangleq [y_1(n), \ldots, y_{N_r}(n)]^T$, $s(n) \triangleq [s_1(n), \ldots, s_{N_t}(n)]^T$,
- $H \triangleq [h_1, \ldots, h_{N_r}]^T$, $h_k \triangleq [h_{k,1}, \ldots, h_{k,N_t}]^T$, for $k = 1, \ldots, N_r$,
- $\Theta^r(n) \triangleq \text{diag}(e^{j\theta_1^r(n)}, \ldots, e^{j\theta_{N_r}^r(n)})$ is an $N_r \times N_r$ matrix, $\Theta^t(n) \triangleq \text{diag}(e^{j\theta_1^t(n)}, \ldots, e^{j\theta_{N_t}^t(n)})$ is an $N_t \times N_t$ matrix,
- $B(n) \triangleq [\beta_1(n), \ldots, \beta_{N_r}(n)]^T$ is an $N_r \times N_t$ matrix with $\beta_k(n) \triangleq [\beta_{k,1}, \ldots, \beta_{k,N_t}]^T$,
- $A \triangleq [\alpha_1, \ldots, \alpha_{N_r}]^T$ is the $N_r \times N_t$ channel gain matrix with $\alpha_k \triangleq [\alpha_{k,1}, \ldots, \alpha_{k,N_t}]^T$; and
- $w(n) \triangleq [w_1(n), \ldots, w_{N_r}(n)]^T$.

The following remarks are in order:

**Remark 1:** Given that both the channel gains, $A$, and phase fluctuations, $B(n)$, in (4) are assumed to be unknown, the TS at the beginning of each frame is used to jointly estimate a total of $2N_rN_t$ 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 $N_rN_t$ phase noise parameters, $B(n)$, 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 $N_rN_t$ parameters $\beta_{k,\ell}(n)$, $\forall k, \ell$, and the multi-parameter BCRLB are very difficult. 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, $\Delta^{[t]}_\ell(n)$ and $\Delta^{[r]}_k(n)$, $\forall k, \ell$, in to the additive noise term, $w(n)$ and transform the joint estimation of channel gains, $A$, and phase fluctuations, $B(n)$ 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 Data-Aided Estimation (DAE)

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

$$\lambda \triangleq [\lambda_1^T, \ldots, \lambda_{N_r}^T]^T,$$

(5)

where $\lambda_k \triangleq [\alpha_k^T, \beta_k^T(n)]^T$, 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 $n$, $A$ and $B(n)$. 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 Eq. (2) can be rewritten as

$$\theta^{[t]}_\ell(n-L_t) = \theta^{[t]}_\ell(n) - \sum_{m=-L_t+1}^n \Delta^{[t]}_\ell(m),$$

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

(6)

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

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

$$y_k(n-L_t) \approx \sum_{\ell=1}^{N_r} h_{k,\ell}e^{j\beta_{k,\ell}(n)}$$

$$\times \left( 1 - j \sum_{m=-L_t+1}^n \Delta^{[r]}_k(m) + \Delta^{[t]}_\ell(m) \right)$$

$$\times s_\ell(n-L_t) + w_k(n-L_t),$$

(8)

since for practical oscillators the phase noise innovation variances are small [11], [12] and the Taylor series expansion of the term, $e^{-j\Delta^{[r]}_k(m) + \Delta^{[t]}_\ell(m)}$, for small phase innovations $\Delta^{[t]}_\ell$ and $\Delta^{[r]}_k$ can be approximated by

$$y_k(n-L_t) = \sum_{\ell=1}^{N_r} h_{k,\ell}e^{j\beta_{k,\ell}(n)}$$

$$\times \left( 1 - j \sum_{m=-L_t+1}^n \Delta^{[r]}_k(m) + \Delta^{[t]}_\ell(m) \right)$$

$$\times s_\ell(n-L_t) + w_k(n-L_t),$$

(7a)

$$\approx \sum_{\ell=1}^{N_r} \alpha_{k,\ell}e^{j\beta_{k,\ell}(n)}$$

$$\times \left( 1 - j \sum_{m=-L_t+1}^n \Delta^{[r]}_k(m) + \Delta^{[t]}_\ell(m) \right)$$

$$\times s_\ell(n-L_t) + w_k(n-L_t).$$

(7b)
\[ e^{-j\left(\sum_{m=n-L+1}^{n}(\Delta_k^m(m) + \Delta_{k,\ell}^m(m))\right)} \approx 1 - \frac{j}{\sigma_N^2} \sum_{m=n-L+1}^{n} \left(\Delta_k^m(m) + \Delta_{k,\ell}^m(m)\right). \]  

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., \( \sigma_N^2 = 10^{-2} \text{ rad}^2 \) [11], [12], [19], [27], [29].

The received training signal at the \( k \)'th antenna in (7) can be written in vector form as

\[ y_k \approx \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} s_\ell \]
\[ -j \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} s_\ell \otimes \psi_{k,\ell} + w_k, \]

where

- \( y_k \triangleq [y_k(n-L_l+1), \ldots, y_k(n)]^T \),
- \( s_\ell \triangleq [s_\ell(n-L_l+1), \ldots, s_\ell(n)]^T \),
- \( w_k \triangleq [w_k(n-L_l+1), \ldots, w_k(n)]^T \), and
- \( \psi_{k,\ell} \triangleq \sum_{m=n-L_l+1}^{n} \Delta_k^m(m) + \Delta_{k,\ell}^m(m), \sum_{m=n-L_l+1}^{n} \Delta_k^m(m) + \Delta_{k,\ell}^m(m), \ldots, 0]^T \).

The received signals at all \( N_t \) antennas, \( y \), can be written as \( y = [y_1^T, \ldots, y_{N_t}^T]^T \). Note that based on Assumption A2, \( \Delta_k^m(n), \Delta_{k,\ell}^m(n), \Delta_{k,\ell}^m(n), \) and \( \Delta_{k,\ell}^m(n), \forall k \neq \tilde{k}, \ell \neq \tilde{\ell}, n \neq \tilde{m}, \) are mutually independent. In addition, it follows from the assumptions in Section II and (1) that \( w_k(n) \) and \( w_k(m), \forall k \neq \tilde{k}, n \neq \tilde{m}, \) are mutually independent. Subsequently, the received signal vector, \( y \), is distributed as \( \mathcal{CN}(\mu_y, \Sigma_y) \), where the mean and the \( N_rL_t \times N_rL_t \) covariance matrix of \( y \), \( \mu_y \) and \( \Sigma_y \), respectively, are given by

\[ \mu_y \triangleq \begin{bmatrix} \mu_{y,1}^T \cdots \mu_{y,N_r}^T \end{bmatrix}^T, \]
\[ \Sigma_y \triangleq \begin{bmatrix} \Sigma_{y,1,1} & \cdots & \Sigma_{y,1,N_r} \\ \vdots & \ddots & \vdots \\ \Sigma_{y,N_r,1} & \cdots & \Sigma_{y,N_r,N_r} \end{bmatrix}. \]

In (11), \( \mu_{y,\ell,\ell} \) is given by \( \mu_{y,\ell,\ell} = \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j\beta_{k,\ell}(n)} s_\ell \) and as shown in Appendix A, the \( L_t \times L_t \) sub-matrices \( \Sigma_{y,\ell,\ell} \), for \( \ell, \tilde{\ell} = 1, \ldots, N_t \) and \( k, \tilde{k} = 1, \ldots, N_r \), can be determined as

\[ \text{FIM}_{k,\ell} = \begin{bmatrix} 2R \left\{ U_k^H \Sigma_{y,1}^{-1} U_k \right\} + \Pi_{11}(k,\tilde{k}) & -23 \left\{ U_k^H \Sigma_{y,1}^{-1} U_k Y_k \right\} + \Pi_{12}(k,\tilde{k}) \\ 23 \left\{ Y_k U_k^H \Sigma_{y,1}^{-1} Y_k \right\} + \Pi_{21}(k,\tilde{k}) & 2R \left\{ Y_k U_k^H \Sigma_{y,1}^{-1} U_k Y_k \right\} + \Pi_{22}(k,\tilde{k}) \end{bmatrix}, \]
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 $N_rN_r \times N_tN_r$, FIM for estimation of overall channels in (4), $P(n)$, FIM $^{[P(n)]}$ can be also determined in a similar fashion as above, where the $N_r \times N_t$ submatrices $\text{FIM}_{k,\bar{k}}^{[P(n)]}$, for $k, \bar{k}=1, \cdots, N_r$, are given by

$$\text{FIM}_{k,\bar{k}}^{[P(n)]} = \left[ 2R \left\{ \left( U_k^{[P(n)]} \right)^H \Sigma_y^{-1} U_k^{[P(n)]} \right\} + \Pi^{[P(n)]}_{11}(k, \bar{k}) \right]. \quad (16)$$

In (16), $U_k^{[P(n)]} \triangleq \left[ 0_{(k-1)L_r \times N_t}, S^T, 0_{(N_r-k)L_r \times N_t} \right]^T$, $\Pi^{[P(n)]}_{11}(k, \bar{k}) \triangleq \text{Tr} \left\{ \Sigma_y^{-1} \partial_{\sigma_y} y_{k,\bar{k}} \Sigma_y^{-1} \partial_{\sigma_y} y_{k,\bar{k}} \right\}$, $\bar{y}_{k,\bar{k}}$ denotes the $k$th row and $\bar{k}$th column element of $P(n)$ and $\partial_{\sigma_y} y_{k,\bar{k}}$, for $m = 1, \cdots, N_r$, is determined as

$$\frac{\partial y_{k,\bar{k}}}{\partial y_{m,p}(n)} = \begin{cases} 0, & m = \bar{k} \neq k \\ \frac{p_{m,p}^H(n) s_l \circ \sigma_\Delta \Psi + \sigma_\Delta^2 \Psi}{\sum_{p \neq l}^{N_r} \left( p_{m,p}^H(n) s_l \circ \sigma_\Delta \Psi \right)}, & m = k = \bar{k} \\ \frac{p_{\bar{k},\bar{l}}^H(n) s_l \circ \sigma_\Delta \Psi}{\sum_{\bar{l} \neq \bar{k}}^{N_r} \left( p_{\bar{k},\bar{l}}^H(n) s_l \circ \sigma_\Delta \Psi \right)}, & m = \bar{k} \neq k \\ 0, & m = k \neq \bar{k} \neq k \\ 0, & m \neq k; m \neq \bar{k} \neq k \end{cases} \quad (17)$$

Using (15), the FIM in (17) can be applied to find the CRLB for estimation of the overall MIMO channels including phase noise, $P(n)$. Note that as shown in Section V-B, $P(n)$ is used at the MIMO receiver to equalize the effect of channels and phase noise. Therefore, the CRLB on the estimation accuracy of $P(n)$ 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., $\sigma_\Delta^2 = 10^{-5}$ 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 $n$th symbol’s phase noise parameters, $B(n)$, 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, $\bar{\lambda}$, is given by

$$\bar{\lambda} = [\beta^T_{1}(n), \cdots, \beta^T_{N_r}(n)]^T. \quad (18)$$

The $N_rN_r \times N_tN_r$, FIM for the DDE, $\text{FIM}_{k,\bar{k}}$, can be derived in a similar fashion as that of DAE in (13), where the $N_r \times N_t$ submatrices of $\text{FIM}$ denoted by $\text{FIM}_{k,\bar{k}}^{[P(n)]}$, for $k, \bar{k}=1, \cdots, N_r$, are given by

$$\text{FIM}_{k,\bar{k}}^{[P(n)]} = 2R \left\{ Y_{k,\bar{k}} \Sigma_y^{-1} Y_{k,\bar{k}} + \tilde{N}_{22}(k, \bar{k}) \right\}. \quad (19)$$

$^6$Similar steps as the ones outlined in Section III-A and Appendix A are used and, therefore, are not repeated here.
In (19), \( \overline{\Sigma} \) denotes the covariance matrix of the observation sequence in the decision-directed scenario, \( \overline{y}_k \triangleq [y_k(n - L_o), \ldots , y_k(n - 1)]^T \), which is determined by replacing \( \Psi \) with \( \overline{\Psi} \) in (12), where \( \overline{\Psi} \) is given by

\[
\overline{\Psi} \triangleq \begin{bmatrix}
L_o & L_o - 1 & L_o - 2 & \cdots & 1 \\
L_o - 1 & L_o - 1 & L_o - 2 & \cdots & 1 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1 & 1 & \cdots & 1
\end{bmatrix}.
\]

In addition, in (19), \( \overline{\Phi}_k \) is an \( N_k L_o \times N_t \) matrix, \( \overline{S} \triangleq [\overline{s}_1, \ldots , \overline{s}_{N_t}]^T \), and \( \overline{\Pi}_{22}(k, k) \) is determined by replacing \( \Sigma_y \) with \( \overline{\Sigma}_y \) in \( \overline{\Pi}_{22}(k, k) \) below (14). The CRLB for the decision directed case, \( \text{CRLB}(\lambda) = \text{diag}(\text{FTM}^{-1}) \).

**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 and an EKF for decision-directed tracking of phase noise parameters over a frame are derived.

**A. Data-Aided Estimation (DAE)**

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

\[
p(y | A, B(n)) = \frac{1}{(2\pi)^{N_s} | \Sigma_y |} \exp \left\{ - \frac{1}{2} (y - \mu_y)^H \Sigma_y^{-1} (y - \mu_y) \right\}.
\]

Based on (20) and since the covariance matrix, \( \Sigma_y \) 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 \( A \) and \( B(n) \), \( \hat{A} \) and \( \hat{B}^{\text{DAE}}(n) \), respectively, can be determined as

\[
\hat{A}, \hat{B}^{\text{DAE}}(n) = \arg\min_{A, B} (y - \mu_y)^H (y - \mu_y).
\]

The cost function in (21) can be modified as

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

where \( Y \triangleq [y_1, \ldots , y_{N_t}]^T \) and \( S \) and \( P(n) \) are defined in (14) and (4), respectively. Using well-known methods, the LS estimate of \( P(n) \), \( \hat{P}(n) \), can be shown to be

\[
\hat{P}(n) = Y S^H (S S^H)^{-1} = \frac{Y S^H}{L_t},
\]

where the second equality in (23) follows from Assumption A1, i.e., \( S S^H = L_t I_{N_t \times N_t} \). As demonstrated in [11], [12] and discussed in Section II, for most practical oscillators, the phase noise variances, \( \sigma_{\Delta^2}^2 \) and \( \sigma_{\Delta^2}^2 \), \( \forall k, \ell \) 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 parameters are determined by

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

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

\[
\log \left( p(y | \hat{A}, B(n)) \right) = \log (\det(\overline{\Sigma}_y)) + (y - \overline{\mu}_y)^H \overline{\Sigma}_y^{-1} (y - \overline{\mu}_y) - \log L_t - N_t - 1,
\]

where \( \overline{\mu}_y = \sum_{n=1}^{N_t} \alpha_n e^{j \beta_n} \bar{s}_n \), and \( y, \bar{\Sigma}_y, \bar{s}_n \) are defined in Section III-B. Using (25), decision-directed WLS estimates of the \( n \)th symbol’s \( N_t N_r \) phase noises, \( B^{\text{DAE}}(n) \), is given by

\[
\hat{B}^{\text{DAE}}(n) = \arg\min_{B(n)} (y - \mu_y)^H \Sigma_y^{-1} (y - \mu_y).
\]

In order to reduce the computational complexity, the term \( \log (\det(\overline{\Sigma}_y)) \) in (26) is omitted, resulting in a WLS estimator. Numerical results demonstrate that \( \log (\det(\overline{\Sigma}_y)) \) 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 the above initialization, AP converges to the covariance matrix, matrix that can be determined straightforwardly as

\[ Q = \begin{bmatrix} Q_{1,1} & \cdots & Q_{1,N_r} \\ \vdots & \ddots & \vdots \\ Q_{N_r,1} & \cdots & Q_{N_r,N_r} \end{bmatrix} \]

(27)

where \( \Delta_k,\ell(n) = \Delta_k^{[r]}(n) + \Delta_\ell^{[t]}(n) \) is the sum of the phase noise innovations for the \( k \)th transmit and the \( \ell \)th receive antennas. Using (27), the unknown state vector \( \phi(n) \) is given in (24), and \( \hat{\phi}(n) \) is evaluated at \( \phi(n) \), and

\[ \phi(n) = \phi(n-1) + \delta(n), \]

(28)

where the state noise vector \( \delta(n) \) is distributed as \( \delta(n) \sim \mathcal{N}(0,N_r \times 1, \mathbf{Q}) \). The state noise covariance matrix, \( \mathbf{Q} = \mathbb{E}\{\delta(n)\delta(n)^T\} \), is an \( N_r \times N_r \) matrix that can be determined straightforwardly as

\[ Q = \begin{bmatrix} Q_{1,1} & \cdots & Q_{1,N_r} \\ \vdots & \ddots & \vdots \\ Q_{N_r,1} & \cdots & Q_{N_r,N_r} \end{bmatrix} \]

(29)

\[ Q_{k,k} = \begin{cases} \sigma^2_\Delta_{1,\ell}, & k = \ell \\ \sigma^2_\Delta_{k,\ell}, & k \neq \ell \end{cases}, \]

(30)

The observation equation for the Kalman filter is given by

\[ y(n) = z(n) + w(n) = (\mathbf{A} \circ e^{\mathbf{B}(n)}) s(n) + w(n), \]

(31)

where \( z(n) = \{z_1(n), \ldots, z_{N_r}(n)\}^T = (\mathbf{A} \circ e^{\mathbf{B}(n)}) s(n) \). Since the observation equation in (31) is a non-linear function of the unknown state vector \( \phi(n) \), 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 \( z(n) \) with respect to the state vector \( \phi(n) \) as

\[ D(\phi(n)) = \frac{\partial z(n)}{\partial \phi^T(n)} \]

(32)

\[ = \left[ \begin{array}{ccc} \frac{\partial z_1(n)}{\partial \beta_{k,1}(n)} & \cdots & \frac{\partial z_1(n)}{\partial \beta_{N_r,1}(n)} \\ \vdots & \ddots & \vdots \\ \frac{\partial z_{N_r}(n)}{\partial \beta_{k,1}(n)} & \cdots & \frac{\partial z_{N_r}(n)}{\partial \beta_{N_r,1}(n)} \end{array} \right] \]

(33)

where \( D(\phi(n)) \) denotes the \( N_r \times N_r \) matrix evaluated at \( \phi(n) \), and

\[ \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} = \left[ \begin{array}{c} \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} \\ \frac{\partial z_k(n)}{\partial \beta_{k,1}(n)} \\ \vdots \end{array} \right] \]

(34)

\[ = \left[ \begin{array}{c} \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} \\ \frac{\partial z_k(n)}{\partial \beta_{k,1}(n)} \\ \vdots \end{array} \right] \]

(35)

\[ = \left[ \begin{array}{c} \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} \\ \frac{\partial z_k(n)}{\partial \beta_{k,1}(n)} \\ \vdots \end{array} \right] \]

(36)

\[ = \left[ \begin{array}{c} \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} \\ \frac{\partial z_k(n)}{\partial \beta_{k,1}(n)} \\ \vdots \end{array} \right] \]

(37)

\[ = \left[ \begin{array}{c} \frac{\partial z_k(n)}{\partial \beta_{k,\ell}(n)} \\ \frac{\partial z_k(n)}{\partial \beta_{k,1}(n)} \\ \vdots \end{array} \right] \]

(38)

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

\[ 7 \text{In this paper } \chi \text{ is generally selected to be a small value, e.g., } \chi = 0.1, \text{ since as shown in Section V-B the data-aided estimates are accurate.} \]
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 by $C_{\text{WLS}} = C_{\text{WLS}[M]} + C_{\text{WLS}[A]}$. The notations $C_{\text{WLS}[M]}$ and $C_{\text{WLS}[A]}$ denote the number of complex multiplications and additions, respectively, used by the WLS estimator.

In order to quantitatively compare the computational complexity of the proposed WLS and EKF estimator, we have evaluated $C_{\text{WLS}}$ and $C_{\text{EKF}}[M]$ for a $2 \times 2$ MIMO system. We assumed $L_o = N_t$ and in order to ensure that the performance of the proposed WLS is close to the CRLB, the step size, $\kappa = 10^{-2}$ and $N = 4$ 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 $\sigma_{\Delta[k]}^2 = \sigma_{\Delta[k]}^2 = \sigma_{\Delta[k]}^2$, $\forall k, t$, and $\sigma_o^2 = 1/\text{SNR}$. The MIMO channel matrix is generated as a sum of LoS and non-line-of-sight (NLoS) components such that the overall channel matrix, $H$, is given by [14]

$$H = \sqrt{\frac{K}{K+1}} H_{\text{LoS}} + \sqrt{\frac{1}{K+1}} H_{\text{NLoS}},$$

where $K$ denotes the Rician factor, i.e., the ratio between the power of the LoS and NLoS channel components [62, p.52], $H_{\text{LoS}}$ and $H_{\text{NLoS}}$, respectively. The elements of $H_{\text{LoS}}$ 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, $\varphi_t$ and $\varphi_r$, respectively, are set to $\varphi_t = 5^\circ$ and $\varphi_r = 0^\circ$, the carrier frequency is set to 60 GHz, and unless otherwise specified $K = 2$ dB. The elements of $H_{\text{NLoS}}$ are modeled as independent and identically distributed (i.i.d) complex Gaussian random variables with $CN(0,1)$. Walsh-Hadamard codes with binary phase-shift keying (BPSK) or quadrature phase-shift keying (QPSK) are used for the Ts.

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

\[ C_{\text{WLS}}^{[M]} = N(N_t N_r) 2\pi \frac{N_t}{\kappa} \left\{ N_t [2(N_t^2 + 3) + (N_t^2 - N_r)] [N_t(N_r + 1)] \right\} \]

\[ + \left\{ (N_t L_o) + (N_t L_o)^3 + N_t N_r(1 + L_o) \right\} \]

\[ \text{in (39)} \]

\[ C_{\text{WLS}}^{[A]} = N(N_t N_r) 2\pi \frac{N_t}{\kappa} \left\{ N_t [L_o^2((N_t - 1)^2 + N_t)] + (N_t^2 - N_r) [L_o^2(N_t - 1)] + N_r L_o(N_r L_o) \right\} \]

\[ \text{in (39)} \]
gain and phase noise 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, $P(n)$, and time-varying phases, respectively, versus SNR. The CRLB in (15) is numerically evaluated for different phase noise variances, e.g., $\sigma^2 = [10^{-3}, 10^{-4}, 10^{-5}] \text{rad}^2$. Note that, $\sigma^2 = 10^{-3} \text{rad}^2$, 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, $w_k$, is negligible, the additive noise corresponding to phase noise limits the estimation accuracy. The same noise also limits the estimation accuracy of the $n$th 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 against 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 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 $L_\alpha = 2$. 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, $\vec{y}_k = [y_k(n - L_\alpha), \cdots , y_k(n - 1)]^T$ can be used while estimating the $n$th symbol’s phase noise. Therefore, as observed in Fig. 6 the estimation accuracy of phases at time $n$, $B(n)$, cannot reach below the variance of the white Gaussian phase noise innovations, $\Delta^2(n)$ and $\Delta^2\ell(n)$, $\forall k, \ell$.

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 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

![Fig. 4. MSE for DAE of channel, $P(n)$, for a 2 × 2 MIMO system for different phase noise variances.](image)

![Fig. 5. MSE for DAE of phase noise for a 2 × 2 MIMO system for different phase noise variances.](image)
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

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

is used to equalize the effect of phase noise and channel gains. Note that since the matrices $\Theta^{[r]}(n)$ and $\Theta^{[I]}(n)$ 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., $\sigma^2_\Delta = 10^{-3}$ 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 1.5dB with SNR=20dB and $\sigma^2_\Delta = 10^{-4}$ 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 3dB with SNR=20dB). 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 illustrates the BER performance of a 2 × 2 MIMO system using the proposed EKF phase tracking algorithm with imperfect decision feedback for different pilot spacings. 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 $2 \times 2$ and $4 \times 4$ 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 $4 \times 4$ MIMO system is only 3dB 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 $4 \times 4$ system, at high SNR, overall system performance is dominated by phase noise and compared to a $2 \times 2$ 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 $2 \times 2$ MIMO system using the proposed EKF estimator for different Rician factors $K$. $K = [-3, 2, 10]$ 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 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 $4 \times 4$ 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 of $[0.001, 0.002, 0.004]$ and synchronization overheads of $[10\%, 16.6\%]$. The results in Fig. 14

$^3$The time-varying channels are generated based on the Jakes' spectrum [63].
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 trade-off 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 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., offline 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 A
DERIVATION OF FIM

In this appendix the FIM for the data-aided estimation of $\lambda$ is derived. The received signal vector at the $k$th receive antenna, $y_k \triangleq [y_k(n - L_t + 1), \ldots, y_k(n)]^T$ is given by

$$y_k = \sum_{\ell = 1}^{N_t} \alpha_{k, \ell} e^{j \beta_{k, \ell}(n)} s_\ell - j \sum_{\ell = 1}^{N_t} \alpha_{k, \ell} e^{j \beta_{k, \ell}(n)} s_\ell \otimes \psi_{k, \ell} + w_k,$$  

(1.1)

where

- $\psi_{k, \ell} \triangleq \left[ \sum_{m = n - L_t + 2}^{n} \Delta_k^{(1)}(m) + \Delta_k^{(1)}(m), \ldots, \sum_{m = n - L_t + 3}^{n} \Delta_k^{(1)}(m) + \Delta_k^{(1)}(m), \ldots, 0 \right]^T$,
- $s_\ell \triangleq [s_\ell(n - L_t + 1), \ldots, s_\ell(n)]^T$,
- $w_k \triangleq [w_k(n - L_t + 1), \ldots, w_k(n)]^T$.

Based on the assumptions in Section II, the vector of received signals at all receive antennas, $y \triangleq [y_1^T, \ldots, y_{N_r}^T]^T$, is distributed as $y \sim \mathcal{CN}(\mu_y, \Sigma_y)$, where $\mu_y \triangleq \left[ \mu_{y_1}^T, \ldots, \mu_{y_{N_r}}^T \right]^T$ with $\mu_{y_k} = \sum_{\ell = 1}^{N_t} \alpha_{k, \ell} e^{j \beta_{k, \ell}(n)} s_\ell$, and the $L_t \times L_t$ submatrices $\Sigma_{y_{k, k}}$ of the $N_r L_t \times N_r L_t$ covariance matrix $\Sigma_y$.
matrix, $\Sigma_y$, for $k, \bar{k} = 1, \ldots, N_r$, can be determined as

$$\Sigma_{y,k,\bar{k}} = \mathbb{E}\left[ -j \sum_{\ell=1}^{N_t} \alpha_{k,\ell} e^{j \beta_{k,\ell}(n)} s_{\ell} \otimes \psi_{k,\ell} + w_k \right] (A.2a)$$

$$\times \left( -j \sum_{\ell=1}^{N_t} \alpha_\ell e^{j \beta_{\ell}(n)} s_{\ell} \otimes \psi_{\ell} + w_\ell \right)^H \right] \right),
$$

$$\sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{k,\ell} e^{j \beta_{k,\ell}(n)} e^{-j \beta_{k,\ell}(n)}
\times \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{\ell} e^{j \beta_{\ell}(n)} e^{-j \beta_{\ell}(n)}
\times \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{k,\ell} e^{j \beta_{k,\ell}(n)} e^{-j \beta_{k,\ell}(n)}
\times \sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{\ell} e^{j \beta_{\ell}(n)} e^{-j \beta_{\ell}(n)}$$

$$\times (s_{k} \otimes \psi_{k}) (s_{\bar{k}} \otimes \psi_{\bar{k}})^H \right), \quad k \neq \bar{k}$$

(A.2b)

$$\sum_{\ell=1}^{N_t} \sum_{\bar{\ell}=1}^{N_t} \alpha_{k,\ell} e^{j \beta_{k,\ell}(n)} e^{-j \beta_{k,\ell}(n)}
\times (s_{k} \otimes \psi_{k}) (s_{\bar{k}} \otimes \psi_{\bar{k}})^H \right), \quad k \neq \bar{k}$$

(A.2c)

Note that (A.2b) follows from (A.2a) due to the assumption that the AWGN and phase noise innovations $\Delta^t[k]$, $\Delta^r[k]$, $\forall k, \ell$, 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, $y$, has a complex Gaussian distribution, the $i$th row and $i$th column entry of the submatrix $\text{FIM}_{k,\bar{k}}$ for $\ell, \bar{\ell} = 1, \ldots, 2N_t$ and $k, \bar{k} = 1, \ldots, N_r$ is given by [61]

$$\text{FIM}_{k,\bar{k}}{\ell,\bar{\ell}} = 2\Re \left\{ \frac{\partial \mu_y}{\partial \alpha_{k,\ell}} \frac{\partial \mu_y}{\partial \lambda_{k,\ell}} \right\} + \text{Tr} \left\{ \frac{\partial \Sigma_y}{\partial \lambda_{k,\ell}} \frac{\partial \Sigma_y}{\partial \lambda_{\bar{k},\bar{\ell}}} \right\},$$

(A.3)

where $\lambda_{k,\ell}$ denotes the $\ell$th element of the vector of parameters corresponding to the $k$th receive antenna, $\lambda_k$ in (A.1). In order to evaluate the FIM, the derivatives $\frac{\partial \mu_y}{\partial \alpha_{k,\ell}}$ and $\frac{\partial \mu_y}{\partial \beta_{k,\ell}(n)}$ need to be determined, where

$$\frac{\partial \mu_y}{\partial \alpha_{k,\ell}} = \left[ 0, \ldots, 0, e^{j \beta_{k,\ell}(n)} s_{\ell}^T, 0, \ldots, 0 \right]^T$$

$$\frac{\partial \mu_y}{\partial \beta_{k,\ell}(n)} = \left[ 0, \ldots, 0, \alpha_{k,\ell} e^{j \beta_{k,\ell}(n)} s_{\ell}^T, 0, \ldots, 0 \right]^T$$

(A.4)

In addition, the derivatives of the sub-matrices $\Sigma_{y,k,\bar{k}}$ in (A.2), $\frac{\partial \Sigma_{y,k,\bar{k}}}{\partial \alpha_{m,\ell}}$ and $\frac{\partial \Sigma_{y,k,\bar{k}}}{\partial \beta_{m,\ell}(n)}$, for $m = 1, \ldots, N_r$, can be evaluated as shown in (A.5) and (A.6) at the bottom of this page.

After substituting the derivatives in (A.4)-(A.6) into (A.3) and carrying out straightforward algebraic manipulations, the

$$10 \Delta^t_k(n)$$ and $\Delta^r_k(m)$ are assumed to be mutually independent for $m \neq n$. 

elements of $\text{FIM}_{k,\ell}$ can be obtained as

$$\text{FIM}_{k,\ell} = 2Re \left\{ \frac{\partial \mu_y^H}{\partial \alpha_{k,\ell}} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \mu_y}{\partial \beta_k} \right\} + \text{Tr} \left( \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \alpha_{k,\ell}} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \alpha_{k,\ell}} \right) \right\}$$  \hspace{1cm} (A.7a)

$$\text{FIM}_{k,\ell+1} = 2Re \left\{ \frac{\partial \mu_y^H}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \mu_y}{\partial \beta_k} \right\} + \text{Tr} \left( \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \right) \right\}$$  \hspace{1cm} (A.7b)

$$\text{FIM}_{k,\ell+1} = 2Re \left\{ \frac{\partial \mu_y^H}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \mu_y}{\partial \beta_k} \right\} + \text{Tr} \left( \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \right) \right\}$$  \hspace{1cm} (A.7c)

$$\text{FIM}_{k,\ell+1} = 2Re \left\{ \frac{\partial \mu_y^H}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \mu_y}{\partial \beta_k} \right\} + \text{Tr} \left( \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \sum_{q=1}^{\gamma} \frac{1}{\beta_k} \frac{\partial \Sigma_y}{\partial \beta_k} \right) \right\}$$  \hspace{1cm} (A.7d)

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

REFERENCES


[64] Ali A. Nasir received his B.Sc. (first-class honors) degree in Electrical Engineering from the University of Engineering and Technology (UET), Lahore, Pakistan in 2007. He is currently pursuing his Ph.D. degree from the Research School of Engineering, the Australian National University, Canberra, Australia. 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 at Chalmers University of Technology, Gothenburg, Sweden for six months from April till October 2011. Ali 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. His research interests are in the area of synchronization and channel estimation in cooperative communication, MIMO and OFDM systems.

[65] Steven D. Blostein (S’83-M’88-SM’96) received his 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 from the University of Illinois, Urbana-Champaign, in 1985 and 1988, respectively. He has been on the faculty in the Department of Electrical and Computer Engineering Queen’s University since 1988 and currently holds the position of Professor. From 2004-2009 he was Department Head. From 1999-2003, 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 in Ottawa. 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. He 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 has been serving on technical program committees for IEEE Communications Society conferences that include ICC, Globecom and WCN. He has been serving as an editor of IEEE Transactions on Wireless Communications since 2007. He is a registered professional engineer in Ontario and a Senior Member of IEEE.
Thomas Eriksson received the B.Sc. (1st class honours) degree in Electrical Engineering from the University of Engineering & Technology, Lahore, Pakistan in 2000. He received the PhD degree in Electrical Engineering from the University of Queensland, Brisbane, Australia in Dec. 2004. Since March 2005, he has been a Lecturer in the Research School of Engineering, College of Engineering & Computer Science, The Australian National University, Canberra, Australia. His current research interests are in wireless communications and signal processing, including synchronization in cooperative communication systems, connectivity of ad-hoc networks and vehicular networks and signal processing on the unit sphere. He serves as a Technical Program Committee Member for international conferences such as ICC ’12, PIMRC ’11 and AusCTW ’12. He was a recipient of an International Postgraduate Research Scholarship from the Australian Commonwealth during 2001-2004. He was awarded a University Gold Medal during his undergraduate studies. He has 47 publications to date in refereed international journals and conferences. He is a Member of Institution of Engineers, Australia and a Senior Member of IEEE.

George K. Karagiannidis (S’86-M’88-SM’01-F’05) received the B.E. degree from the University of New South Wales, Sydney, Australia, the M.E. degree from the University of Newcastle, and the Ph.D. degree from the Australian National University, Canberra. He is currently a Professor in the Research School of Engineering, Australian National University. He is a Fellow of the IEEE. His research interests include digital signal processing, digital and wireless communications, and acoustical signal processing.

Tommy Svensson (S’86-M’88-SM’01-F’05) received the B.E. degree from the University of New South Wales, Sydney, Australia, the M.E. degree from the University of Newcastle, and the Ph.D. degree from the Australian National University, Canberra. He is currently a Professor in the Research School of Engineering, Australian National University. He is a Fellow of the IEEE. His research interests include digital signal processing, digital and wireless communications, and acoustical signal processing.