



## Phase noise estimation in OFDM systems

Downloaded from: <https://research.chalmers.se>, 2025-04-19 12:28 UTC

Citation for the original published paper (version of record):

Gävert, B., Coldrey, M., Eriksson, T. (2024). Phase noise estimation in OFDM systems. IEEE Transactions on Communications, In Press. <http://dx.doi.org/10.1109/TCOMM.2024.3493814>

N.B. When citing this work, cite the original published paper.

© 2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, or reuse of any copyrighted component of this work in other works.

# Phase noise estimation in OFDM systems

Björn Gävert, Mikael Coldrey and Thomas Eriksson *Member, IEEE*

**Abstract**—In this paper, we present a comprehensive foundation for the theory and practice of pilot-based carrier phase synchronization of OFDM symbols in the presence of phase noise. The system covers phase noise added in both the transmitter and receiver together with a frequency-selective channel. A novel, low-complexity, phase noise estimation method is presented, and we show that it is close to optimal over a wide dynamic range of phase noise. Additionally, a method for approximating the phase noise over the OFDM symbol using principal component analysis is proposed, where the number of parameters to estimate can be significantly reduced with minimal impact on performance.

**Index Terms**—Wireless communication, OFDM, carrier synchronization, phase noise estimation, frequency domain pilots.

## I. INTRODUCTION

CARRIER phase noise is a well-known, quite often dominating, problem in wireless and optical communication systems. The growing demand for high capacity typically drives the need for improved spectral efficiency, which drives system sensitivity and associated requirements. Moreover, with the recent interest in radio systems operating at higher carrier frequencies, still maintaining high spectral efficiency, the design of oscillators has become more challenging since the phase noise typically gets worse as the frequency increases. Synchronization aspects is an implementation challenge of beyond-5G and 6G communication [1]. Hence, the phase noise requirements are continuously increasing and methods for estimating and suppressing the effects of phase noise are of high importance.

The phase noise residue, or phase noise degradation, can be reduced by either utilizing better performing oscillators or by implementing better performing phase estimation algorithms. Improved oscillator performance typically leads to higher power consumption and increased cost (e.g., more complex hardware) as indicated in [2], and better performing phase noise estimation algorithms quite often lead to increased complexity and, in many cases, to increased system delay. Hence, the balance between the complexity aspects of the system and reasonable degradation of the system performance must often be treated with great consideration.

Orthogonal frequency division multiplexing (OFDM) is a popular method for designing signals in wireless systems intended for dispersive channel conditions, and is therefore successfully used in e.g., 4G/5G and Wi-Fi systems. However, the OFDM signal is sensitive to non-linear distortion and carrier phase noise. The phase noise introduces phase rotation

on each OFDM subcarrier, or a common phase error (CPE), and a crosstalk between the OFDM subcarriers, or intercarrier interference (ICI). The ICI comes from the fact that the phase noise rotates the time domain version of the OFDM signal, while the information is mapped in the frequency domain representation. It is well known that a time domain multiplication becomes a convolution in the frequency domain, see, for example, [3, Section 6.3]. Hence, the Fourier transform of the phase noise convolves with the frequency domain representation of the OFDM signal (or filters the OFDM signal in the frequency domain).

Estimation of phase noise in wireless and optical OFDM systems has been thoroughly studied. Even though empirical methods are proposed, the phase noise estimation commonly utilizes hard (or soft) decisions of the payload, often in an iterative manner. A method using payload decisions is typically referred to as a decision directed method. For reducing problems related to decision errors (for example, error propagation), pilot symbols, e.g., known symbols can be mapped on certain known subcarriers. The pilots are often combined with unreliable symbol decisions when performing phase estimation. A time domain phase locked loop (PLL) approach is presented in [4], and a quasi-pilot-assisted system is proposed in [5]. A blind phase estimator is presented in [6]. A pilot-based CPE estimator, which is enhanced by decision directed iterations, is analyzed in [7]. An iterative decision directed approach to estimate the minimum mean square error (MMSE), modeling phase noise using a power series basis, is proposed in [8]. MMSE estimation of phase noise using a subset of the discrete Fourier transform (DFT) base (for reducing complexity) and decisions after a channel decoder is analyzed in [9], and similar approaches can be seen in [10] and [11]. An iterative, decision directed, phase noise and channel estimation utilizing the geometrical properties of the phase noise is presented in [12]. Few publications consider phase noise in both the transmitter and receiver, however, an interesting joint estimation of transmitter and receiver phase noise has been analyzed in [13]. By linearizing the frequency domain expressions, an approximate linear system is proposed, and the phase noise is estimated using MMSE with decisions using a channel decoder.

Since decision directed methods in general suffer from high complexity and error propagation, pilot-based phase noise estimation has also been studied in literature. Frequency-consecutive pilots for performing pilot-based phase estimation on a wider bandwidth have been proposed in many papers, e.g., [14] and [15]. An interesting system is studied in [16], which uses frames of OFDM symbols, forming a packet. The first OFDM symbol, containing only pilots, is used to estimate the channel in the presence of phase noise. The phase noise estimation of the following, payload dominated, OFDM

This research has been carried out in joint research project financed by Stiftelsen för Strategisk Forskning (SSF) and Ericsson.

B. Gävert and M. Coldrey are with Ericsson AB, Gothenburg, Sweden.

T. Eriksson is with the Department of Electrical Engineering, Chalmers University of Technology, Gothenburg, Sweden.

symbols is then based on forming piece-wise continuous functions using pilot-based CPE estimates of consecutive OFDM symbols (interpolation). A phase noise estimation method using scattered pilots is proposed in [17], where phase noise modeled using the discrete cosine transform (DCT) base is compared with a DFT base modeling, and estimation is based on least squares (LS) minimizing the error of the pilots with respect to the phase noise. A powerful method for estimating very strong phase noise using scattered pilots is presented in [18]. As in [12], the method approximates the phase noise in the frequency domain using a geometry-preserving model. The estimation is performed using constrained LS and performance is claimed to be optimal. However, due to the constraints, the method becomes complex. In addition, a normalized LS (NLS) is proposed. The complexity is significantly reduced, however, so is also the performance.

In this paper, we extend the studies to complement earlier work on phase estimation based on a small number of pilots in OFDM systems affected by phase noise. Major contributions can be summarized as:

- 1) A novel method, with significantly lower complexity than other comparable methods in literature, is presented, which, under realistic conditions, is close to optimal in an MMSE sense when estimating phase noise in OFDM systems using distributed pilots allocated in the frequency domain. A complexity reducing method is proposed, which approximates the phase noise using principal component analysis (PCA), where the number of estimated parameters can be significantly reduced with minor impact on accuracy of the estimated phase noise.
- 2) We demonstrate the importance of utilizing phase noise statistics when estimating the phase noise of an OFDM system using only a small number of distributed pilots.
- 3) We present the optimal and rate-maximizing number of pilots as a function of phase noise level.
- 4) We analyze a realistic system with phase noise both in the transmitter and receiver and a Rayleigh block fading channel, and present a joint estimation of all phase noise. We, furthermore, demonstrate that the system can be well approximated as a system with receiver-only phase noise if the channel and the phase noise fulfill certain conditions.

## II. NOTATION

This section contains a description of the notation used in this paper.

Scalar variables are denoted with lowercase letters, e.g., the scalar complex rotation  $e^{i\phi}$ . Vectors are denoted with bold lowercase letters, e.g., the signal vector  $\mathbf{x}$ . If not explicitly stated in the text, the length of a vector is  $N$ , i.e. the length of an OFDM symbol. Matrices are denoted with bold uppercase letters, e.g., the Fourier transform matrix  $\mathbf{F}$ . If not explicitly stated in the text, the size of a matrix is  $N \times N$ .

Vectors and matrices are either random or deterministic, and no special notation is used for either type. Instead, the text related to each equation clearly defines whether a variable is deterministic or random.

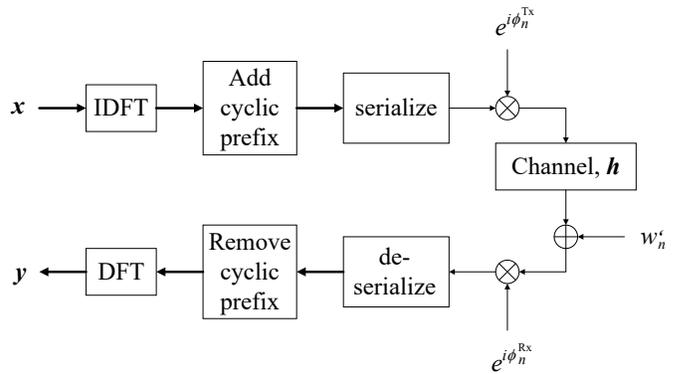


Fig. 1. OFDM system with phase noise in the transmitter and receiver.

## III. SYSTEM DESCRIPTION

This section contains a general description of the system, the phase noise modeling aspects related to the system and the channel model.

### A. System model

An OFDM system, visualized in Figure 1, transmitting an OFDM symbol of length  $N$  over a noisy channel with memory affected by phase noise is considered. The transmitted signal vector consists of elements,  $x_n$ , which is a set of independent QAM symbols of power  $\sigma_x^2$  with a uniform distribution. Among the QAM symbols, there is also a sparse amount of known pilot symbols of power  $\sigma_x^2$  at known locations in the frequency domain, where the number of pilots  $N_{\text{pilot}} \ll N$ . A cyclic prefix (CP) is appended to the transmitted time domain signal, where the length of the CP corresponds to the memory of the channel. The phase noise at the transmitter is  $\phi_n^{\text{Tx}}$ , and the phase noise at the receiver is  $\phi_n^{\text{Rx}}$ , both represented as phasors,  $e^{i\phi_n^{\text{Tx}}}$  and  $e^{i\phi_n^{\text{Rx}}}$ , which phase rotates the signal in the transmitter and receiver. The AWGN,  $w'_n$ , is white and  $\mathcal{CN}(0, 1/N\sigma_w^2)$ , and independent of the QAM symbols and phase noise. The channel filtering is represented by  $\mathbf{h}$ , and models a Rayleigh block fading channel with an impulse response which is in this paper truncated to 10% of the OFDM symbol, i.e., the length of the CP is  $N/10$ . The complex and normal distributed filter taps of  $\mathbf{h}$  have a square average amplitude that decays exponentially.

The received frequency domain signal,  $\mathbf{y}$  in Figure 1, can be described in matrix form, where the phase noise rotates the time domain signal, i.e., an element-wise phase rotation of the time domain samples of the signal in the transmitter and receiver. Therefore, the received OFDM symbol can be formulated in matrix form as

$$\begin{aligned} \mathbf{y} &= \mathbf{F} e^{i\Phi_{\text{Rx}}} \left( \mathbf{F}^{-1} \mathbf{H} \mathbf{F} e^{i\Phi_{\text{Tx}}} \mathbf{F}^{-1} \mathbf{x} + \mathbf{e}' + \mathbf{w}' \right) \\ &= \underbrace{\mathbf{F} e^{i\Phi_{\text{Rx}}} \mathbf{F}^{-1}}_{\text{time domain}} \mathbf{H} \underbrace{\mathbf{F} e^{i\Phi_{\text{Tx}}} \mathbf{F}^{-1}}_{\text{time domain}} \mathbf{x} + \mathbf{e} + \mathbf{w}, \end{aligned} \quad (1)$$

where the AWGN,  $w$ , is  $\mathcal{CN}(\mathbf{0}, \sigma_w^2 \mathbf{I})$ . The matrix  $\mathbf{F}$  is the discrete Fourier transform (DFT) defined as

$$\mathbf{F} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & e^{-i2\pi/N} & \cdots & e^{-i2\pi(N-1)/N} \\ 1 & e^{-i2\pi 2/N} & \cdots & e^{-i2\pi 2(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-i2\pi(N-1)/N} & \cdots & e^{-i2\pi(N-1)(N-1)/N} \end{pmatrix}. \quad (2)$$

The matrix  $\mathbf{F}^{-1} = 1/N \mathbf{F}^H$ , is the inverse discrete Fourier transform (IDFT). Due to the CP, the cyclic filtering can be modeled as a diagonal channel matrix,  $\mathbf{H} = \text{diag}\{\mathbf{F}\mathbf{h}\}$ , containing the DFT of the channel impulse response  $\mathbf{h}$ , where  $\text{diag}\{\mathbf{a}\}$  is a diagonal matrix containing the vector  $\mathbf{a}$ . In other words,  $\mathbf{H}$  contains the frequency domain channel gains for each OFDM subcarrier. The model assumes that the average channel gain is one (the mean square of the diagonal of  $\mathbf{H}$  is one), which corresponds to a transmit output power control of the complete signal. The vector  $\mathbf{e}$ , which is a term that corrects for the modeling error related to the cyclic filtering of the transmitter phase noise, is described further down. The phase noise matrix,  $\Phi$ , is diagonal and contains the phase noise process vector variable, and  $e^{i\Phi}$  is defined as a diagonal element-wise exponential resulting in appropriate phasor matrices. The transmitter phase noise is independent of the receiver phase noise, and both of them are independent with the QAM symbols and the AWGN.

The cyclic prefix correction term,  $\mathbf{e}$ , is included to achieve a correct matrix model that corresponds to Figure 1. The correction term,  $\mathbf{e}$ , can be modeled using a diagonal  $N \times N$  matrix  $\mathbf{D}$ , where the diagonal is 0 for the non-prefix data and 1 for the prefix data. Hence, the data used for the CP is the nonzero last part of the length  $N$  vector  $\mathbf{D}\mathbf{F}^{-1}\mathbf{x}$ . The cyclic filtering modeling error related to the transmitter phase noise can be formulated as

$$\begin{aligned} \mathbf{e} &= -\mathbf{F}e^{i\Phi_{\text{Rx}}}(\mathbf{I} - \mathbf{D})\mathbf{F}^{-1}\mathbf{H}\mathbf{F}\mathbf{D}e^{i\Phi_{\text{Tx}}}\mathbf{F}^{-1}\mathbf{x} \\ &\quad + \mathbf{F}e^{i\Phi_{\text{Rx}}}(\mathbf{I} - \mathbf{D})\mathbf{F}^{-1}\mathbf{H}\mathbf{F}\mathbf{D}e^{i\tilde{\Phi}_{\text{Tx}}}\mathbf{F}^{-1}\mathbf{x} \\ &= \mathbf{F}e^{i\Phi_{\text{Rx}}}(\mathbf{I} - \mathbf{D})\mathbf{F}^{-1}\mathbf{H}\mathbf{F}\mathbf{D}\left(e^{i\tilde{\Phi}_{\text{Tx}}} - e^{i\Phi_{\text{Tx}}}\right)\mathbf{F}^{-1}\mathbf{x}, \quad (3) \end{aligned}$$

where the phase noise matrix  $\tilde{\Phi}_{\text{Tx}}$  is the transmitter phase noise from the previous  $N$  time domain samples including the CP. This formulation can be derived directly by analyzing the effects of the cyclic filtering. It can be understood as removing the cyclic part related to the transmitter phase noise<sup>1</sup> of the current symbol, and replacing it with the corresponding part of the actual CP with the correct phase noise. In reality, the error related to  $\mathbf{e}$  is often insignificant and will only affect performance for extreme transmitter phase noise in a combination with very selective channels.

### B. Channel

The modeling assumes that the frequency domain channel,  $\mathbf{H}$ , is known with some statistical uncertainty, or channel estimation error. The OFDM symbols are often organized in

frames, consisting of several OFDM symbols grouped together in time, where an initial OFDM symbol contains a substantial amount of pilots (known QAM symbols) to estimate the channel in the presence of phase noise and other distortion, see, for example, [16]. The following OFDM symbols in a frame contain significantly less pilots and, instead, more payload in the form of unknown QAM symbols, and the channel knowledge is utilized in the receiver under the assumption that the channel state is constant over the frame. It is these payload-dominated OFDM symbols, having a sparse amount of pilots, that are considered in this paper.

The channel estimate is denoted

$$\hat{\mathbf{H}} = \mathbf{H} + \Delta_{\mathbf{H}}, \quad (4)$$

and the estimation error,  $\Delta_{\mathbf{H}}$ , is a complex zero mean normal distributed variable which is uncorrelated with the signal, the AWGN and the phase noise. The diagonal covariance matrix of the channel estimation error is

$$\mathbf{C}_{\Delta H} = \sigma_{\Delta H}^2 \mathbf{I}. \quad (5)$$

### C. Phase noise

Both oscillator and phase noise modeling have been thoroughly studied in the literature. A simple and often valid model is the Wiener process, which captures the basic and important properties of phase noise, i.e., a Gaussian random process with linearly time-increasing variance [19]. The modeling of the Wiener process as a discrete time process is analyzed in [20], as phase noise in reality is a continuous time process. The Wiener process has, due to its inherent simplicity, often been used in different theoretical phase noise studies. Also in this paper, the time-discrete Wiener process will be used to model transmitter and receiver phase noise. Hence, the phase noise is formulated as

$$\phi_n = \phi_{n-1} + \Delta_n, \quad (6)$$

where the innovation of the transmitter phase noise is white and  $\mathcal{N}(0, \sigma_{\Delta_{\text{Tx}}}^2)$ , and the innovation of receiver phase noise is white and  $\mathcal{N}(0, \sigma_{\Delta_{\text{Rx}}}^2)$ . Over a batch of  $N$  phase noise samples, the start phase,  $\phi_0$ , of the transmitter and receiver is uniform<sup>2</sup>  $[0 \ 2\pi]$ .

The two phase noise matrices in (1),  $\Phi_{\text{Tx}}$  and  $\Phi_{\text{Rx}}$ , can be separated into four different random parts; a zero average transmitter phase noise,  $\Theta_{\text{Tx}}$ , with corresponding random average,  $\bar{\phi}^{\text{Tx}}$ , and a zero average receiver phase noise,  $\Theta_{\text{Rx}}$ , with corresponding random average  $\bar{\phi}^{\text{Rx}}$ . Since the start phase of the phase noise process vectors,  $\phi_0$ , is uniform  $[0 \ 2\pi]$ ,

<sup>2</sup>The first OFDM symbol in a frame, Section III-B, will have a uniform start phase due to a new realization of the channel. From a theoretical perspective, the phase estimation could utilize the last estimated phase of the previous OFDM symbols within a frame. However, this is not considered as the possible improvement is small for moderate to high phase noise levels.

<sup>1</sup>The part of the OFDM symbol which is used for creating a CP does not have the same phase noise as the actual CP.

TABLE I  
VISUALIZATION OF THE DIFFERENT STEPS WHEN FORMULATING THE ESTIMATOR.

- Step 1: Taylor expand phasor matrices
$e^{i\Theta} = \sum_{k=0}^{\infty} \frac{(i\Theta)^k}{k!}$
- Step 2: Formulate phase noise as vectors
$i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} = i\mathbf{F}\text{diag}\{\boldsymbol{\theta}_{\text{Rx}}\}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} = i\mathbf{F}\text{diag}\{\mathbf{F}^{-1}\mathbf{H}\mathbf{x}\}\boldsymbol{\theta}_{\text{Rx}}$
- Step 3: Resolve rank deficiency due to the unknown payload
$\mathbf{x}_n \rightarrow \tilde{\mathbf{x}}_n$
- Step 4: Resolve phase offset ambiguity
$\boldsymbol{\theta}_n^{\text{Tx}}, \boldsymbol{\theta}_n^{\text{Rx}} \rightarrow \boldsymbol{\theta}_n^{\text{Tx}}, \boldsymbol{\theta}_n^{\text{Tx}} + \boldsymbol{\theta}_n^{\text{Rx}}$
- Step 5: Formulate estimator
$e^{i\bar{\theta}}, \bar{\boldsymbol{\theta}}^{\text{Tx}}, \bar{\boldsymbol{\theta}}^{\text{Rx}}$

the average phase,  $\bar{\phi} \bmod 2\pi = 1/N \sum_k \phi_k \bmod 2\pi$ , is uniform  $[0, 2\pi]$ . Therefore, (1) can be formulated as

$$\begin{aligned}
\mathbf{y} &= \mathbf{F} e^{i\bar{\phi}^{\text{Rx}}} e^{-i\bar{\phi}^{\text{Rx}}} e^{i\Phi_{\text{Rx}}} \mathbf{F}^{-1} \\
&\quad \times \mathbf{H} \mathbf{F} e^{i\bar{\phi}^{\text{Tx}}} e^{-i\bar{\phi}^{\text{Tx}}} e^{i\Phi_{\text{Tx}}} \mathbf{F}^{-1} \mathbf{x} + \mathbf{e} + \mathbf{w} \\
&= \underbrace{e^{i\bar{\phi}^{\text{Rx}}} e^{i\bar{\phi}^{\text{Tx}}}}_{e^{i\bar{\theta}}} \mathbf{F} e^{-i\bar{\phi}^{\text{Rx}}} e^{i\Phi_{\text{Rx}}} \mathbf{F}^{-1} \\
&\quad \times \mathbf{H} \mathbf{F} \underbrace{e^{-i\bar{\phi}^{\text{Tx}}} e^{i\Phi_{\text{Tx}}}}_{e^{i\Theta_{\text{Tx}}}} \mathbf{F}^{-1} \mathbf{x} + \mathbf{e} + \mathbf{w} \\
&= \underbrace{e^{i\bar{\theta}}}_{\text{CPE}} \mathbf{F} e^{i\Theta_{\text{Rx}}} \mathbf{F}^{-1} \mathbf{H} \mathbf{F} e^{i\Theta_{\text{Tx}}} \mathbf{F}^{-1} \mathbf{x} + \mathbf{e} + \mathbf{w}. \quad (7)
\end{aligned}$$

The two random averages,  $\bar{\phi}^{\text{Tx}} + \bar{\phi}^{\text{Rx}}$ , have been combined into a common CPE phasor,  $e^{i\bar{\theta}}$ , where  $\bar{\theta} \bmod 2\pi$  is uniform  $[0, 2\pi]$ . The resulting model, (7), will show important when formulating an estimator.

The covariance matrices of the CPE-normalized phase noise,  $\mathbf{C}_{\theta}^{\text{Tx}}$  and  $\mathbf{C}_{\theta}^{\text{Rx}}$  are derived in Appendix A, where the diagonals of  $\Theta_{\text{Tx}}$  and  $\Theta_{\text{Rx}}$  are  $\boldsymbol{\theta}_{\text{Tx}}$  and  $\boldsymbol{\theta}_{\text{Rx}}$  in vector form. The elements of  $\mathbf{C}_{\theta}$  are

$$\begin{aligned}
\mathbf{C}_{\theta, k, l} &= \max(l, k) \sigma_{\Delta}^2 + \sigma_{\Delta}^2 \frac{l^2 - l(2N + 1)}{2N} \\
&\quad + \sigma_{\Delta}^2 \frac{k^2 - k(2N + 1)}{2N} + \sigma_{\Delta}^2 \frac{2N^2 + 3N + 1}{6N}. \quad (8)
\end{aligned}$$

To reduce complexity, an approximation of the phase noise is presented in Appendix B. Similar to PCA, see, for example, [21], the phase noise can be approximated as

$$\boldsymbol{\theta} \approx \mathbf{U}_{\theta}^k \tilde{\boldsymbol{\theta}}_k, \quad (9)$$

where the  $N \times k$  matrix  $\mathbf{U}_{\theta}^k$  contains the  $k$  singular vectors corresponding to the  $k$  strongest singular values of  $\mathbf{C}_{\theta}$  and  $\tilde{\boldsymbol{\theta}}_k$  corresponds to  $k$  phase noise parameters. The covariance matrix of  $\tilde{\boldsymbol{\theta}}_k$  can be formulated as

$$\mathbf{C}_{\tilde{\theta}} = \mathbf{U}_{\theta}^{k\text{T}} \mathbf{C}_{\theta} \mathbf{U}_{\theta}^k, \quad (10)$$

which is a  $k \times k$  diagonal matrix.

#### IV. PILOT-BASED PHASE NOISE ESTIMATION

This section contains a complete analysis and the modeling aspects to find an MMSE estimate of the phase noise in

the system model, (7). A reformulation of the system model is presented in Section IV-A, a solution to rank deficiency related to sparse known pilots embedded in random payload is proposed in Section IV-B, and the estimator is presented in Section IV-C. The different steps used in formulating the estimator are visualized in Table I, and the corresponding steps are marked with bold font in the text.

The intention is to formulate a close to optimal estimator of the phase noise processes in (7), based on a small number of known pilot symbols without requirements on how they are allocated within the OFDM symbol. For configurations with few pilots, it is important that the estimator utilizes the known statistics of the phase noise, and the known statistics of the AWGN and other noise sources. This normally boils down to formulating the minimum mean square error (MMSE) estimate of the phase noise.

##### A. System model reformulation

In this section, a system model reformulation is derived for finding a close-to-optimal pilot-based phase estimator using standard textbook MMSE estimators. By reformulating (7), estimation of the phase can be made simpler.

**Step 1:** Looking at (7), the received input signal is a nonlinear function of the phase noise. Optimal estimation is a complicated nonlinear problem, and the resulting estimator will be of high complexity. Instead, it is of interest to formulate a less complex linear estimator, which is close to optimal. A close to optimal linear estimator can be formulated by several reformulations of (7). As a first step, the phasors are Taylor expanded, resulting in

$$\begin{aligned}
\mathbf{y} &= e^{i\bar{\theta}} \mathbf{F} \left( \underbrace{\sum_{k=0}^{\infty} \frac{(i\Theta_{\text{Rx}})^k}{k!}}_{e^{i\Theta_{\text{Rx}}}} \right) \\
&\quad \times \mathbf{F}^{-1} \mathbf{H} \mathbf{F} \left( \underbrace{\sum_{k=0}^{\infty} \frac{(i\Theta_{\text{Tx}})^k}{k!}}_{e^{i\Theta_{\text{Tx}}}} \right) \mathbf{F}^{-1} \mathbf{x} + \mathbf{e} + \mathbf{w} \\
&= e^{i\bar{\theta}} \mathbf{F} \left( \mathbf{I} + \sum_{k=1}^{\infty} \frac{(i\Theta_{\text{Rx}})^k}{k!} \right) \mathbf{F}^{-1} \mathbf{H} \mathbf{F} \\
&\quad \times \left( \mathbf{I} + \sum_{k=1}^{\infty} \frac{(i\Theta_{\text{Tx}})^k}{k!} \right) \mathbf{F}^{-1} \mathbf{x} + \mathbf{e} + \mathbf{w}, \quad (11)
\end{aligned}$$

where the zero order terms have been separated from the sums. After multiplication, the resulting phase noise degraded signal part contains four terms. Separating also the first element of the sum from the two terms containing a single sum, the higher order terms of the Taylor expansion can be collected in a single vector  $\mathbf{n}$ , and (11) can be formulated as

$$\begin{aligned}
\mathbf{y} &= e^{i\bar{\theta}} \mathbf{H} \mathbf{x} + i e^{i\bar{\theta}} \mathbf{F} \Theta_{\text{Rx}} \mathbf{F}^{-1} \mathbf{H} \mathbf{x} + i e^{i\bar{\theta}} \mathbf{H} \mathbf{F} \Theta_{\text{Tx}} \mathbf{F}^{-1} \mathbf{x} \\
&\quad + \mathbf{n} + \mathbf{e} + \mathbf{w}, \quad (12)
\end{aligned}$$

where

$$\begin{aligned} \mathbf{n} = & e^{i\bar{\theta}} \mathbf{F} \sum_{k=2}^{\infty} \frac{(i\Theta_{\text{Rx}})^k}{k!} \mathbf{F}^{-1} \mathbf{H} \mathbf{x} \\ & + e^{i\bar{\theta}} \mathbf{H} \mathbf{F} \sum_{k=2}^{\infty} \frac{(i\Theta_{\text{Tx}})^k}{k!} \mathbf{F}^{-1} \mathbf{x} \\ & + e^{i\bar{\theta}} \mathbf{F} \sum_{k=1}^{\infty} \frac{(i\Theta_{\text{Rx}})^k}{k!} \mathbf{F}^{-1} \mathbf{H} \mathbf{F} \sum_{k=1}^{\infty} \frac{(i\Theta_{\text{Tx}})^k}{k!} \mathbf{F}^{-1} \mathbf{x}. \end{aligned} \quad (13)$$

It should be noted that (12) can be seen as linear with respect to the CPE phasor,  $e^{i\bar{\theta}}$ , and a rotated, complex, CPE-normalized phase noise,  $ie^{i\bar{\theta}}\Theta_{\text{Tx}}$  and  $ie^{i\bar{\theta}}\Theta_{\text{Rx}}$ , and the nonlinear phase noise part is captured by  $\mathbf{n}$ . The linear form is an important step in order to formulate a linear MMSE estimator, see, for example, [22, Eq. 12.26].

The phase noise is normally not allowed to be a completely dominating source of degradation in many systems. For example, wireless access systems designed for transporting QAM symbols with high orders of modulation. The linear form of (12) is valid as long as the covariance of  $\mathbf{n}$  is small enough, i.e., the phase noise is small enough. Numerical evaluation of (13) shows that the induced norm of the covariance matrix of  $\mathbf{n}$  (since the covariance matrix is  $N \times N$ , this corresponds to the highest eigenvalue) is insignificant compared to the corresponding norm of the AWGN covariance matrix for low to moderate phase noise levels at high SNR. Furthermore, the impact of  $\mathbf{n}$  is analyzed in Section VI.

**Step 2:** For cases where diagonal matrices multiply vectors, the diagonal of the diagonal matrix can be exchanged with the following vector, i.e.,  $\text{diag}\{\mathbf{a}\}\mathbf{b} = \text{diag}\{\mathbf{b}\}\mathbf{a}$ . In order to have the unknown phase noise as vectors on the right side in the corresponding terms of (12), the diagonal of the diagonal phase noise matrices,  $\Theta_{\text{Tx}}$  and  $\Theta_{\text{Rx}}$ , can be exchanged with the following vectors, for example,

$$\begin{aligned} i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} &= i\mathbf{F}\text{diag}\{\boldsymbol{\theta}_{\text{Rx}}\}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} \\ &= i\mathbf{F}\underbrace{\text{diag}\{\mathbf{F}^{-1}\mathbf{H}\mathbf{x}\}}_{\mathbf{R}_{Hx}}\boldsymbol{\theta}_{\text{Rx}} \\ &= i\mathbf{F}\mathbf{R}_{Hx}\boldsymbol{\theta}_{\text{Rx}}. \end{aligned} \quad (14)$$

Here  $\mathbf{R}_{Hx}$  is diagonal and contains the time domain signal  $\mathbf{F}^{-1}\mathbf{H}\mathbf{x}$ , and the vector  $\boldsymbol{\theta}_{\text{Rx}}$  is the diagonal of  $\Theta_{\text{Rx}}$ . The term  $i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x}$  can be reformulated in a similar manner as  $i\mathbf{H}\mathbf{F}\mathbf{R}_x\boldsymbol{\theta}_{\text{Tx}}$ , where  $\mathbf{R}_x$  is diagonal and contains the time domain signal  $\mathbf{F}^{-1}\mathbf{x}$ . Therefore, (12) can be written as

$$\begin{aligned} \mathbf{y} = & e^{i\bar{\theta}} \mathbf{H} \mathbf{x} + ie^{i\bar{\theta}} (\mathbf{F}\mathbf{R}_{Hx} \mathbf{H}\mathbf{F}\mathbf{R}_x) \begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ & + \mathbf{n} + \mathbf{e} + \mathbf{w}, \end{aligned} \quad (15)$$

where the transmitter and receiver phase noise have been collected in one vector of length  $2N$ .

Estimating the phase noise in (15) is not straightforward since  $\mathbf{R}_{Hx}$  and  $\mathbf{R}_x$  mostly contain unknown payload symbols. The problem relates to the fact that the deterministic pilot part of  $\mathbf{R}_{Hx}$  and  $\mathbf{R}_x$  is rank deficient, with a rank  $N_{\text{pilot}} \ll N$ . A popular approach to handle the unknown payload is to estimate the phase noise using a decision directed method, either with

or without including a channel code, see, for example, [9]. This typically comes with significant complexity and problems with error propagation. Pilot-based MMSE estimates of the phase noise over a larger bandwidth have been demonstrated. However, only for systems where the pilots are adjacent in frequency<sup>3</sup>, see, for example, [14]. As shown in [17], the phase noise can also be estimated with LS, minimizing a cost function based on distributed pilots with respect to the phase noise, but with a significant performance penalty since the phase noise statistics are omitted.

## B. Solving rank deficiency

This section presents a method for treating the random payload part in the matrices  $\mathbf{R}_{Hx}$  and  $\mathbf{R}_x$  in (15), which allows for any allocation of pilots within the OFDM symbol, scattered or consecutive, used for estimating the phase noise.

**Step 3:** To handle aspects related to the multiplication of the unknown phase noise and the random payload in (15), estimates of the random matrices  $\mathbf{R}_{Hx}$  and  $\mathbf{R}_x$  are introduced. This is solved here by using an estimate (described below) of the entire pilot/data vector  $\mathbf{x}$ , denoted  $\hat{\mathbf{x}}$ , resulting in  $\mathbf{R}_{H\hat{x}}$  and  $\mathbf{R}_{\hat{x}}$ . Hence, by adding and subtracting the estimated parts, (15) can be formulated as

$$\begin{aligned} \mathbf{y} = & e^{i\bar{\theta}} \mathbf{H} \mathbf{x} + ie^{i\bar{\theta}} (\mathbf{F}\mathbf{R}_{Hx} \mathbf{H}\mathbf{F}\mathbf{R}_x) \begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ & + ie^{i\bar{\theta}} (\mathbf{F}(\mathbf{R}_{H\hat{x}} - \mathbf{R}_{Hx}) \mathbf{H}\mathbf{F}(\mathbf{R}_{\hat{x}} - \mathbf{R}_x)) \begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ & + \mathbf{n} + \mathbf{e} + \mathbf{w} \\ = & e^{i\bar{\theta}} \mathbf{H} \mathbf{x} + ie^{i\bar{\theta}} (\mathbf{F}\mathbf{R}_{H\hat{x}} \mathbf{H}\mathbf{F}\mathbf{R}_{\hat{x}}) \begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ & + \underbrace{ie^{i\bar{\theta}} (\mathbf{F}(\mathbf{R}_{Hx} - \mathbf{R}_{H\hat{x}}) \mathbf{H}\mathbf{F}(\mathbf{R}_x - \mathbf{R}_{\hat{x}}))}_{\mathbf{r}} \begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ & + \mathbf{n} + \mathbf{e} + \mathbf{w}. \end{aligned} \quad (16)$$

When introducing the known full-rank matrices,  $\mathbf{R}_{H\hat{x}}$  and  $\mathbf{R}_{\hat{x}}$ , in (16), the phase noise is mapped without losing any degrees of freedom, but an error,  $\mathbf{r}$ , appears as a consequence. It should be noted that the reformulated model, (16), and the original model, (7), are equal.

Since the phase noise has been left unfiltered (transformed without losing degrees of freedom), the distribution of pilots becomes irrelevant using (16). The estimation of the CPE phasor,  $e^{i\bar{\theta}}$ ,  $\boldsymbol{\theta}_{\text{Tx}}$  and  $\boldsymbol{\theta}_{\text{Rx}}$  can be performed using any rows, or equations, of (16). However, since most frequencies of the received signal,  $\mathbf{y}$ , are completely dominated by the unknown payload, only the pilot frequencies can essentially be used. An  $N_{\text{pilot}} \times N$  matrix,  $\mathbf{T}$ , is introduced to map the  $N$  dimensional vectors to the pilot space, i.e.,  $\mathbf{T}\mathbf{x}$  will result in a vector

<sup>3</sup>The phase noise of the received OFDM signal is the convolution of the Fourier transform of the phase noise and the payload/pilot symbols. If the pilots are not adjacent, most of the phase noise will drown in the payload.

of length  $N_{\text{pilot}}$  containing only pilots,  $\mathbf{x}_{\text{pilot}}$ . From (16), the resulting pilot-space observations become

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &\triangleq \mathbf{T}\mathbf{y} \\ &= e^{i\bar{\theta}}\mathbf{T}\mathbf{H}\mathbf{x} + ie^{i\bar{\theta}}\mathbf{T}\left(\mathbf{F}\mathbf{R}_{H\hat{x}}\mathbf{H}\mathbf{F}\mathbf{R}_{\hat{x}}\right)\begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ &\quad + \mathbf{T}\mathbf{n} + \mathbf{T}\mathbf{r} + \mathbf{T}\mathbf{e} + \mathbf{T}\mathbf{w}, \end{aligned} \quad (17)$$

where the subscript *pilot* indicates that only the pilot frequencies are included.

Noting that the phase noise,  $\boldsymbol{\theta}_{\text{Tx}}$  and  $\boldsymbol{\theta}_{\text{Rx}}$ , and the AWGN have zero mean in (15), it is straightforward to form an estimate  $\hat{\mathbf{x}}$  from the received signal,  $\mathbf{y}$ , as

$$e^{i\bar{\theta}}\hat{\mathbf{x}} = \hat{\mathbf{H}}^{-1}\mathbf{y}, \quad (18)$$

using the channel estimate (4). Since  $\mathbb{E}[\hat{\mathbf{H}}^{-1}\mathbf{y}|\mathbf{x}, e^{i\bar{\theta}}] = e^{i\bar{\theta}}\mathbf{x} + \mathbf{H}^{-1}\mathbb{E}[\mathbf{n}|\mathbf{x}, e^{i\bar{\theta}}] \approx e^{i\bar{\theta}}\mathbf{x}$ , it is approximately an unbiased estimate as long as the higher even order Taylor expansion terms in (13) are small. It is possible to replace the estimated (rotated) pilots, which are a part of  $e^{i\bar{\theta}}\hat{\mathbf{x}}$  in (18), with the corresponding known pilots, but this requires a separate estimate of  $e^{i\bar{\theta}}$  since it is unknown. Instead, the received signal  $\mathbf{y}$ , as in (18), is used as part of a phase noise estimate.

Using the reformulated model, (17), a pilot-based MMSE estimator can be formulated to estimate the phase noise. Such an MMSE estimator will rely on an estimate  $\hat{\mathbf{x}}$ , the statistics of the modeling error  $\mathbf{r} + \mathbf{n}$ , the statistics of  $\mathbf{w}$ , and the statistics of the phase noise, (8).

### C. Phase noise estimation

In this section, an MMSE estimator of the phase noise in (17) is presented using the proposed transmit symbol estimates, (18), and the estimated channel, (4). A reformulation of the phase noise parameters is also introduced to avoid the estimation ambiguity related to having phase noise in both the transmitter and receiver. The reformulation also reveals that the complete phase noise of the system can be approximated as receiver-only if the phase noise is not too dominant and the channel conditions are favorable, i.e., channel is not too "selective".

The observed pilot positions of the received signal, (17), can be formulated using the proposed estimate of the transmitted symbols, (18), resulting in

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &= e^{i\bar{\theta}}\mathbf{T}\mathbf{H}\mathbf{x} + \mathbf{T}\left(\mathbf{F}\mathbf{R}_y\mathbf{H}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\right)\begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ &\quad + \mathbf{T}\mathbf{n} + \mathbf{T}\mathbf{r} + \mathbf{T}\mathbf{e} + \mathbf{T}\mathbf{w}. \end{aligned} \quad (19)$$

Introducing also the channel estimate, (4), gives

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &= e^{i\bar{\theta}}\mathbf{T}(\hat{\mathbf{H}} - \boldsymbol{\Delta}_H)\mathbf{x} \\ &\quad + \mathbf{T}\left(\mathbf{F}\mathbf{R}_y(\hat{\mathbf{H}} - \boldsymbol{\Delta}_H)\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\right)\begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ &\quad + \mathbf{T}\mathbf{n} + \mathbf{T}\mathbf{r} + \mathbf{T}\mathbf{e} + \mathbf{T}\mathbf{w} \\ &= e^{i\bar{\theta}}\mathbf{T}\hat{\mathbf{H}}\mathbf{x} + i\mathbf{T}\left(\mathbf{F}\mathbf{R}_y\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\right)\begin{pmatrix} \boldsymbol{\theta}_{\text{Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix} \\ &\quad + \mathbf{T}\underbrace{\left(\mathbf{r} + \mathbf{n} - \boldsymbol{\Delta}_H(\mathbf{x} + \mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\boldsymbol{\theta}_{\text{Tx}})\right)}_{\mathbf{w}_{\text{tot}}} + \mathbf{T}\mathbf{e} + \mathbf{T}\mathbf{w}, \end{aligned} \quad (20)$$

and a noise vector  $\mathbf{w}_{\text{tot}}$  is introduced to obtain a more compact formulation.

**Step 4:** From equation (7), it is evident that any constant phase rotation of the receiver can be canceled by the corresponding inverse rotation of the transmitter. To illustrate the phase offset ambiguity, (20) can be formulated as

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &= e^{i\bar{\theta}}\mathbf{T}\hat{\mathbf{H}}\mathbf{x} + i\mathbf{T}\mathbf{F}\mathbf{R}_y\underbrace{(\boldsymbol{\theta}_{\text{Rx}} + \delta\mathbf{1})}_{\boldsymbol{\alpha}_{\text{Rx}}} \\ &\quad + i\mathbf{T}\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\underbrace{\left(\boldsymbol{\theta}_{\text{Tx}} - \delta\left(\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\right)^{-1}\mathbf{F}\mathbf{R}_y\mathbf{1}\right)}_{\boldsymbol{\alpha}_{\text{Tx}}} \\ &\quad + \mathbf{T}\mathbf{w}_{\text{tot}} \end{aligned} \quad (21)$$

In this illustrative example, an offset  $\delta$  has been added to the receiver phase noise and subtracted as part of the transmitter phase noise. Hence, for some given realization of the phase noise, there will be an infinite number of possible estimates<sup>4</sup>. Since  $\boldsymbol{\alpha}_{\text{Tx}} + \boldsymbol{\alpha}_{\text{Rx}}$  corresponds to  $\boldsymbol{\theta}_{\text{Tx}} + \boldsymbol{\theta}_{\text{Rx}}$ , any offset uncertainty is canceled in the sum of the transmitter and receiver phase noises. To avoid ambiguity, (20) can be formulated as

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &= e^{i\bar{\theta}}\mathbf{T}\hat{\mathbf{H}}\mathbf{x} + i\mathbf{T}\mathbf{F}\mathbf{R}_y\boldsymbol{\theta}_{\text{Rx}} \\ &\quad + i\mathbf{T}\mathbf{F}\mathbf{R}_y(\boldsymbol{\theta}_{\text{Tx}} - \boldsymbol{\theta}_{\text{Tx}}) \\ &\quad + i\mathbf{T}\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y}\boldsymbol{\theta}_{\text{Tx}} + \mathbf{T}\mathbf{w}_{\text{tot}} \\ &= e^{i\bar{\theta}}\mathbf{T}\hat{\mathbf{H}}\mathbf{x} + i\mathbf{T}\mathbf{F}\mathbf{R}_y\underbrace{(\boldsymbol{\theta}_{\text{Rx}} + \boldsymbol{\theta}_{\text{Tx}})}_{\boldsymbol{\theta}_{\text{Tx+Rx}}} \\ &\quad + i\mathbf{T}\underbrace{\left(\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y} - \mathbf{F}\mathbf{R}_y\right)}_{\boldsymbol{\epsilon}_{\text{Tx}}}\boldsymbol{\theta}_{\text{Tx}} + \mathbf{T}\mathbf{w}_{\text{tot}}, \end{aligned} \quad (22)$$

which represents the phase noise as the sum of the transmitter and receiver phase noise, or system phase noise,  $\boldsymbol{\theta}_{\text{Tx+Rx}}$ , and a residual transmitter phase noise  $\boldsymbol{\theta}_{\text{Tx}}$ .

By looking at (22), it can be understood that the part denoted  $\boldsymbol{\epsilon}_{\text{Tx}}$  becomes  $\mathbf{0}$  if  $\hat{\mathbf{H}} = \mathbf{I}$ . Thus, if the channel is not "too selective" and the phase noise is not too dominant, the total phase noise of the system can be approximated as a receiver-only phase noise. In other words, if the covariance of  $\boldsymbol{\epsilon}_{\text{Tx}}$  is small compared to the noise floor, i.e.,  $\mathbb{E}[\boldsymbol{\epsilon}_{\text{Tx}}^H\boldsymbol{\epsilon}_{\text{Tx}}] \ll \mathbb{E}[\mathbf{w}_{\text{tot}}^H\mathbf{w}_{\text{tot}}]$ , then (22) can be approximated as

$$\mathbf{y}_{\text{pilot}} \approx e^{i\bar{\theta}}\mathbf{T}\hat{\mathbf{H}}\mathbf{x} + i\mathbf{T}\mathbf{F}\mathbf{R}_y\boldsymbol{\theta}_{\text{Tx+Rx}} + \mathbf{T}\mathbf{w}_{\text{tot}}. \quad (23)$$

<sup>4</sup>Note that  $\mathbf{F}\mathbf{R}_y\mathbf{1} = \hat{\mathbf{H}}\mathbf{F}\mathbf{F}^{-1}\hat{\mathbf{H}}^{-1}\mathbf{y} = (\hat{\mathbf{H}}\mathbf{F}\mathbf{R}_{\hat{H}^{-1}y})\mathbf{1}$ .

On the contrary, if this condition is not true, the phase noise must be modeled according to (22).

**Step 5:** Equation (22) is linear with respect to all unknown parameters, where the CPE phasor is the only complex parameter. However, it can be expressed using Cartesian representation, resulting in

$$\begin{aligned} \mathbf{y}_{\text{pilot}} &= \left( \underbrace{\cos(\bar{\theta})}_{\alpha} + i \underbrace{\sin(\bar{\theta})}_{\beta} \right) \mathbf{T} \hat{\mathbf{H}} \mathbf{x} + i \mathbf{T} \mathbf{F} \mathbf{R} \mathbf{R}_y \boldsymbol{\theta}_{\text{Tx+Rx}} \\ &\quad + i \mathbf{T} \left( \hat{\mathbf{H}} \mathbf{F} \mathbf{R}_{\hat{\mathbf{H}}^{-1}y} - \mathbf{F} \mathbf{R}_y \right) \boldsymbol{\theta}_{\text{Tx}} + \mathbf{T} \mathbf{w}_{\text{tot}} \\ &= \mathbf{M} \underbrace{\begin{pmatrix} \alpha \\ \beta \\ \boldsymbol{\theta}_{\text{Tx+Rx}} \\ \boldsymbol{\theta}_{\text{Tx}} \end{pmatrix}}_{\boldsymbol{\theta}_{\text{tot}}} + \mathbf{T} \mathbf{w}_{\text{tot}}, \end{aligned} \quad (24)$$

where

$$\mathbf{M} = \mathbf{T} \begin{pmatrix} \hat{\mathbf{H}} \mathbf{x} & i \hat{\mathbf{H}} \mathbf{x} & \mathbf{F} \mathbf{R}_y & \hat{\mathbf{H}} \mathbf{F} \mathbf{R}_{\hat{\mathbf{H}}^{-1}y} - \mathbf{F} \mathbf{R}_y \end{pmatrix}. \quad (25)$$

A new vector,  $\boldsymbol{\theta}_{\text{tot}}$ , of length  $2N+2$  is introduced, collecting all parameters, and an  $N_{\text{pilot}} \times (2N+2)$  matrix  $\mathbf{M}$  is introduced to obtain a more compact formulation. The covariance matrices of  $\boldsymbol{\theta}_{\text{tot}}$ ,  $\mathbf{C}_{\boldsymbol{\theta}_{\text{tot}}}$ , and the different parts of the covariance matrix related to  $\mathbf{w}_{\text{tot}}$ ,  $\mathbf{C}_{\mathbf{w}_{\text{tot}}}$ , are presented in Appendix C.

It should be noted that the first two columns of  $\mathbf{M}$  are just the known transmitted pilots transformed by the corresponding estimated channel matrix gains. The remaining two matrices are, of course, the pilot rows of  $\mathbf{F} \mathbf{R}_y$  and  $\hat{\mathbf{H}} \mathbf{F} \mathbf{R}_{\hat{\mathbf{H}}^{-1}y} - \mathbf{F} \mathbf{R}_y$ . Or, in other words, an *estimate of the mapping of the receiver and transmitter phase noise onto the observed pilot positions*.

Instead of estimating real parameters from complex data, the system of  $N_{\text{pilot}}$  complex equations in (24) can be replaced with corresponding system of  $2N_{\text{pilot}}$  real equations

$$\begin{pmatrix} \Re\{\mathbf{y}_{\text{pilot}}\} \\ \Im\{\mathbf{y}_{\text{pilot}}\} \end{pmatrix} = \underbrace{\begin{pmatrix} \Re\{\mathbf{M}\} \\ \Im\{\mathbf{M}\} \end{pmatrix}}_{\mathbf{M}_{\text{real}}} \boldsymbol{\theta}_{\text{tot}} + \begin{pmatrix} \Re\{\mathbf{T} \mathbf{w}_{\text{tot}}\} \\ \Im\{\mathbf{T} \mathbf{w}_{\text{tot}}\} \end{pmatrix}, \quad (26)$$

and here an additional  $2N_{\text{pilot}} \times (2N+2)$  matrix,  $\mathbf{M}_{\text{real}}$ , is introduced to further simplify the expression. Since the complex noise,  $\mathbf{w}_{\text{tot}}$ , has uncorrelated real and imaginary parts with equal statistical properties, the corresponding  $2N_{\text{pilot}} \times 2N_{\text{pilot}}$  "real" covariance becomes

$$\mathbf{C}_{\mathbf{w}_{\text{tot}}}^{\text{real}} = \begin{pmatrix} \frac{1}{2} \mathbf{T} \mathbf{C}_{\mathbf{w}_{\text{tot}}} \mathbf{T}^{\text{T}} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2} \mathbf{T} \mathbf{C}_{\mathbf{w}_{\text{tot}}} \mathbf{T}^{\text{T}} \end{pmatrix}. \quad (27)$$

The resulting MMSE estimator of  $\boldsymbol{\theta}_{\text{tot}}$  becomes, using [22, Eq. 12.26],

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\text{tot}} &= \mathbf{C}_{\boldsymbol{\theta}_{\text{tot}}} \mathbf{M}_{\text{real}}^{\text{T}} \\ &\quad \times \left( \mathbf{M}_{\text{real}} \mathbf{C}_{\boldsymbol{\theta}_{\text{tot}}} \mathbf{M}_{\text{real}}^{\text{T}} + \mathbf{C}_{\mathbf{w}_{\text{tot}}}^{\text{real}} \right)^{-1} \begin{pmatrix} \Re\{\mathbf{T} \mathbf{y}\} \\ \Im\{\mathbf{T} \mathbf{y}\} \end{pmatrix}. \end{aligned} \quad (28)$$

The estimate of the receiver phase noise is simply  $\hat{\boldsymbol{\theta}}_{\text{Rx}} = \boldsymbol{\theta}_{\text{Tx+Rx}} - \boldsymbol{\theta}_{\text{Tx}}$ . The linear estimate of the CPE, which is the phase argument of the estimated  $\alpha$  and  $\beta$ , is estimated jointly together with  $\boldsymbol{\theta}_{\text{Tx}}$  and  $\boldsymbol{\theta}_{\text{Rx}}$  and all of them combined constitute

the complete estimate of the phase noise. The error of the estimated CPE can be understood from

$$\begin{aligned} \hat{\theta} &= \angle(\hat{\alpha} + i\hat{\beta}) \\ &= \angle(e^{i\bar{\theta}} + \delta_{\alpha} + i\delta_{\beta}) \\ &= \angle(e^{i\bar{\theta}} + e^{i\bar{\theta}}(\delta'_{\alpha} + i\delta'_{\beta})) \\ &= \bar{\theta} + \angle(1 + \delta'_{\alpha} + i\delta'_{\beta}) \\ &= \bar{\theta} + \arctan\left(\frac{\delta'_{\beta}}{1 + \delta'_{\alpha}}\right), \end{aligned} \quad (29)$$

where the estimation error of  $\hat{\alpha} + i\hat{\beta}$  is  $\delta'_{\alpha} + i\delta'_{\beta}$ . When the CPE phasor estimation error is small, only  $\delta'_{\beta}$  will contribute, i.e., when the SNR  $\gg 0$  dB or when using a sufficient number of pilots at low SNR.

## V. ANALYTICAL PERFORMANCE ANALYSIS

This section defines a minimum Bayesian mean square error (BMSE), where the presented BMSE contains the contributions from the CPE and CPE-normalized phase noise processes. The BMSE is formulated for the asymptotic region, where the phase noise approaches zero.

The lower bound BMSE is based on (17) with genie CPE rotated symbol estimates, and including the estimated channel<sup>5</sup>. This corresponds to (24) with  $\hat{\mathbf{x}} = e^{i\bar{\theta}} \mathbf{x}$ , (4), and  $\mathbf{r} = \mathbf{0}$  (since the estimation error is zero). The resulting equation becomes

$$\begin{aligned} \mathbf{y}_{\text{pilot}}^{\hat{\mathbf{x}}=e^{i\bar{\theta}}\mathbf{x}} &= (\alpha + i\beta) \mathbf{T} \mathbf{H} \mathbf{x} + i \mathbf{T} \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Tx+Rx}} \\ &\quad + i \mathbf{T} (\mathbf{H} \mathbf{F} \mathbf{R}_{\hat{\mathbf{x}}} - \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}}) \boldsymbol{\theta}_{\text{Tx}} \\ &\quad + \mathbf{T} (\mathbf{n} + i \boldsymbol{\Delta}_H (\mathbf{x} + \mathbf{F} \mathbf{R}_{\hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Tx}}) \\ &\quad + i \mathbf{F} \mathbf{R}_{\Delta_H \hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Rx}} + \mathbf{e} + \mathbf{w}), \end{aligned} \quad (30)$$

where, compared to (24), the channel estimate is replaced with the channel and the sign of the corresponding error is, therefore, reversed. The superscript  $\hat{\mathbf{x}} = e^{i\bar{\theta}} \mathbf{x}$  indicates that the estimate corresponds to the true transmitted CPE rotated symbols. The noise parts can be collected in

$$\begin{aligned} \tilde{\mathbf{w}} &= \mathbf{n} + i \boldsymbol{\Delta}_H (\mathbf{x} + \mathbf{F} \mathbf{R}_{\hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Tx}}) + i \mathbf{T} \mathbf{F} \mathbf{R}_{\Delta_H \hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Rx}} \\ &\quad + \mathbf{e} + \mathbf{w}. \end{aligned} \quad (31)$$

Stacking all unknown parameters in one vector,  $\boldsymbol{\theta}_{\text{tot}}$  as in (24), results in

$$\begin{aligned} \mathbf{y}_{\text{pilot}}^{\hat{\mathbf{x}}=e^{i\bar{\theta}}\mathbf{x}} &= (\alpha + i\beta) \mathbf{T} \mathbf{H} \mathbf{x} + i \mathbf{T} \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}} \boldsymbol{\theta}_{\text{Tx+Rx}} \\ &\quad + i \mathbf{T} (\mathbf{H} \mathbf{F} \mathbf{R}_{\hat{\mathbf{x}}} - \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}}) \boldsymbol{\theta}_{\text{Tx}} + \mathbf{T} \tilde{\mathbf{w}} \\ &= \mathbf{N} \boldsymbol{\theta}_{\text{tot}} + \mathbf{T} \tilde{\mathbf{w}}, \end{aligned} \quad (32)$$

where the  $N_{\text{pilot}} \times (2N+2)$  matrix  $\mathbf{N}$  is

$$\mathbf{N} = \mathbf{T} \begin{pmatrix} \mathbf{H} \mathbf{x} & i \mathbf{H} \mathbf{x} & \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}} & \mathbf{H} \mathbf{F} \mathbf{R}_{\hat{\mathbf{x}}} - \mathbf{F} \mathbf{R}_{H\hat{\mathbf{x}}} \end{pmatrix}. \quad (33)$$

Similarly to (26), the system of  $N_{\text{pilot}}$  complex equations in (32) can be expressed as  $2N_{\text{pilot}}$  real equations with the

<sup>5</sup>The bound is here including the channel error. As part of the performance evaluation in Section VI, the bound is also evaluated without channel error, i.e.,  $\boldsymbol{\Delta}_H = \mathbf{0}$ .

corresponding  $2N_{\text{pilot}} \times (2+2N)$  matrix  $\mathbf{N}_{\text{real}}$ . The covariance of  $\boldsymbol{\theta}_{\text{tot}}$  is computed in Appendix C.

The BMSE is evaluated in the asymptotic region where the innovation variance of the transmitter and receiver phase noise both approaches zero, i.e., when the total phase noise approaches zero. In this region, the BMSE is a tight bound. As the phase noise approaches zero, the noise  $\tilde{\mathbf{w}}$  in (31) asymptotically becomes

$$\tilde{\mathbf{w}} \rightarrow \mathbf{w} + e^{i\bar{\theta}} \boldsymbol{\Delta}_H \mathbf{x}; \sigma_{\Delta_{\text{Tx}}}^2, \sigma_{\Delta_{\text{Rx}}}^2 \rightarrow 0 \quad (34)$$

The  $(2N+2) \times (2N+2)$  BMSE matrix is, see [22, Eq. 12.28],

$$\mathbf{B}_{\text{MSE}} = \mathbf{C}_{\theta_{\text{tot}}} - \mathbf{C}_{\theta_{\text{tot}}} \mathbf{N}_{\text{real}}^T \left( \mathbf{N}_{\text{real}} \mathbf{C}_{\theta_{\text{tot}}} \mathbf{N}_{\text{real}}^T + \frac{\sigma_x^2 \sigma_{\Delta H}^2 + \sigma_w^2}{2} \mathbf{I}_{2N_{\text{pilot}}} \right)^{-1} \mathbf{N}_{\text{real}} \mathbf{C}_{\theta_{\text{tot}}}, \quad (35)$$

where  $\mathbf{I}_{2N_{\text{pilot}}}$  is a  $2N_{\text{pilot}} \times 2N_{\text{pilot}}$  identity matrix. The AWGN and the channel estimation error are both rotation invariant, independent, complex, and normal distributed, and the variances of their real and imaginary parts are simply half the variance of the complex-valued counterparts. The BMSE of the total phase noise,  $\boldsymbol{\theta}_{\text{Tx+Rx}}$ , corresponds to the diagonal elements 3 to  $N+2$  of (35). As can be seen in (32), the CPE part corresponds to the top left  $2 \times 2$  part of (35).

Since the matrix  $\mathbf{N}_{\text{real}}$  depends on the payload and the CPE, it is of interest to take the mean of the BMSE matrix,  $\mathbb{E}[\mathbf{B}_{\text{MSE}}]$ . However, an analytical mean is difficult to formulate. Instead,  $\mathbb{E}[\mathbf{B}_{\text{MSE}}]$  can be evaluated numerically by averaging (35) for several realizations of  $\hat{\mathbf{x}}$ , i.e., several realizations of the symbols,  $\mathbf{x}$ , and CPE phase,  $\bar{\theta}$ .

When evaluating the mean of  $\mathbf{B}_{\text{MSE}}$  numerically, the CPE part becomes a diagonal  $2 \times 2$  matrix with close to equal elements which corresponds to the estimation error of the complex CPE phasor. The contribution of the CPE to the CPE-normalized phase noise is, according to (29), one of the two diagonal elements. The BMSE of the CPE is added to the BMSE of  $\boldsymbol{\theta}_{\text{Tx+Rx}}$  in order to achieve the complete estimated phase noise BMSE. As mentioned above, the resulting BMSE is asymptotically tight as  $\sigma_{\Delta_{\text{Tx}}}^2, \sigma_{\Delta_{\text{Rx}}}^2 \rightarrow 0$ . For increasing phase noise, the BMSE will be a lower bound as the noise terms related to the phase noise,  $\mathbf{n}$  and  $\mathbf{e}$ , are not present. The BMSE is compared with the simulated performance of the estimator in Section VI.

## VI. NUMERICAL RESULTS

In this section, different aspects of the pilot-based phase noise estimator are analyzed. The simulation setup is presented in Section VI-A, and the simulated results are presented in Section VI-B.

### A. Simulation setup

The system configuration in all simulations corresponds to an OFDM signal with 1024 subcarriers. The modulation order of the payload is chosen to be 1024 QAM and known 4 QAM symbols for the pilots. The actual choice of modulation of the payload is, in fact, not relevant for the results since the

TABLE II  
SYSTEM PARAMETERS WHICH ARE COMMON FOR ALL SIMULATED RESULTS.

Number of subcarriers ( $N$ )	1024
Cyclic prefix length	10%
Payload modulation	1024 QAM
Pilot modulation	4 QAM
Number of phase noise parameters	30
Number of phase noise parameters, genie case	300
Channel estimator error ( $\sigma_{\Delta H}^2$ )	$5 \times 10^{-4}$

payload is not part of the estimation. The pilots are uniformly distributed in the frequency domain. The innovation power of the phase noise is expressed as the sum of the innovation power of the transmitter and the receiver (a total phase noise), and the phase noise is uniformly distributed between the transmitter and the receiver (the same phase noise level in the transmitter and the receiver). The phase noise error is based on the actual and estimated sum of the transmitter and receiver phase noise. The channel time domain coefficients, i.e.,  $\mathbf{h}$  in  $\mathbf{H} = \mathbf{F}\mathbf{h}$  from (1), decay exponentially and the exponent is chosen such that the channel impulse response is small after  $\approx 50\%$  of the CP (CP is 10% of 1024). Hence, the channel impulse response is represented by 100 taps, where at least 50 taps are significant, which corresponds to a strongly frequency-selective channel. All different simulated cases are performed with different realizations of the channel. The channel estimation error covariance, (5), is chosen to be  $5 \times 10^{-4}$ . The system parameters which are common for all simulations are summarized in Table II.

Phase noise estimation is performed using (28) for several different estimator cases. The *Rx only* case considers an estimator assuming phase noise only at the receiver side, and is defined as leaving out the transmit phase noise and modifying the receiver phase noise to reflect the complete (sum) phase noise of the system according to (23). The *Tx/Rx* case covers an estimator that assumes phase noise in both the transmitter and the receiver. The pilot-based genie version of the estimator, denoted *Genie Tx/Rx*, is defined by replacing  $\hat{\mathbf{x}}$  in  $\mathbf{R}_{\hat{\mathbf{x}}}$  and  $\mathbf{R}_{H\hat{\mathbf{x}}}$  in (17) by  $e^{i\bar{\theta}} \mathbf{x}$ . The linear pilot-based genie version, denoted *Genie Tx/Rx, linear model*, is similar to *Genie Tx/Rx*. However, the simulated data does not contain contributions related to approximations and Taylor expansions, and performance is expected to be close to BMSE. The cases *BMSE with channel error* and *BMSE perfect channel knowledge*, are derived using the methodology formulated in Section V, where the case without channel estimation error corresponds to setting  $\boldsymbol{\Delta}_H \mathbf{x} = \mathbf{0}$  in (32).

The analysis of choosing an appropriate number of parameters to approximate the phase noise, according to (9), is not explicitly presented. The genie cases use 200 phase noise parameters, and the normal estimator cases use 30 parameters (which results in a negligible degradation for the chosen phase noise levels).

### B. Simulation results

In this section, the simulated performance of the estimator is analyzed. First, the empirical covariance, (49), is compared

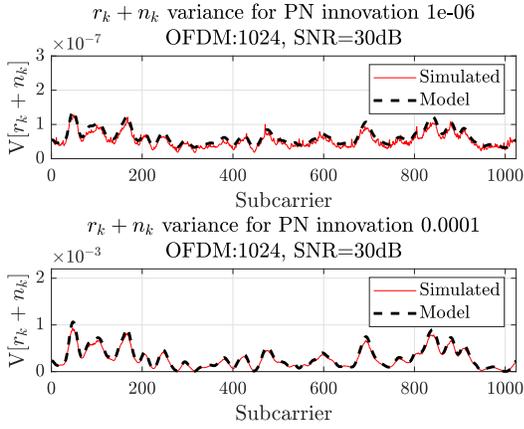


Fig. 2. A comparison between empirical and simulated variance per subcarrier,  $\mathbb{V}[r_k + n_k]$ , of  $r + n$  for two different phase noise levels (innovation variance  $10^{-6}$  and  $10^{-4}$ ). The empirical covariance seems to match the simulated over a large dynamic range of phase noise.

with the corresponding simulated covariance. Second, the simulated mean square error (MSE) is compared with the BMSE according to (35). Third, simulated BER for different cases is analyzed as a function of SNR. Finally, system performance in the form of the average rate per subcarrier is analyzed as a function of phase noise level.

The approximation errors,  $r + n$  in (24), will affect the performance at high phase noise levels, and the empirical covariance, (49), must be verified. The different parameters of (49) are found by simulating (16) and (13) using different system configurations (the model is linear with respect to the different contributing parts). The simulated variance per subcarrier,  $\mathbb{V}[r_k + n_k]$ , is compared with the empirical model for a case with an AWGN SNR of 30 dB and 10 pilots in Figure 2. The simulated data and the empirical model seem to match over a large dynamic range of the phase noise level. The assumption that  $e$  does not contribute is also implicitly verified (since it is included in the simulated data). It should be noted that some parts of (49) can be excluded depending on the SNR, as the channel error and the white noise floor (last two terms) are normally much smaller than the AWGN.

The simulated estimator time domain MSE per OFDM sample for different cases is compared with both *BMSE with channel error* and *BMSE with perfect channel knowledge* in Figure 3 for an AWGN SNR of 40 dB, 10 pilots, and a phase noise innovation power of  $10^{-5}$ . The degradation caused by the channel estimation error is small, but clearly visible. The *Genie Tx/Rx* almost completely overlaps the *Rx only* and *Tx/Rx* cases, indicating that the error related to using an estimate of the transmitted signal, (18), is small. All simulated MSE cases overlap the *BMSE with channel error*. Hence, there is little difference between the different cases for this system configuration, and the most simple, *Rx only*, would be sufficient. This was identified in conjunction with (23), and is of course dependent on the channel.

The estimator in this paper is block-based, which is evident when looking at Figure 3. The MSE close to the edges of the

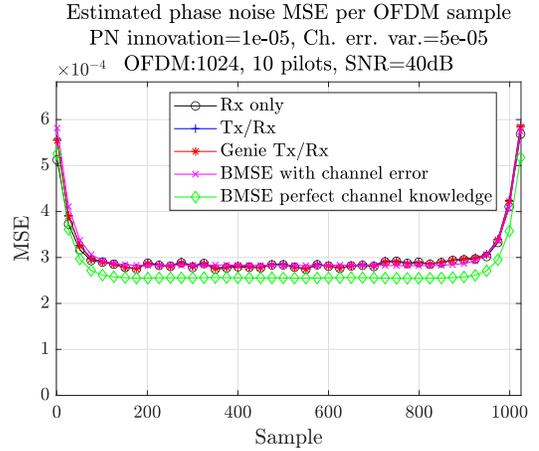


Fig. 3. Simulated time domain MSE per OFDM sample using different estimators, with 10 pilots, compared to BMSE with and without channel estimation error. The channel estimation error variance is in this example  $5 \times 10^{-5}$ . The simulated MSE of *Rx only*, *Tx/Rx* and *Genie Tx/Rx* are almost completely overlapping the *BMSE with channel error*.

OFDM symbol in the time domain is approximately twice the floor, which is a consequence of the fact that only future (or history) is used in the edge estimates. The edge effects are not treated in this paper since they are considered small. In Figure 3, the average increase in MSE due to the edges, i.e., the approximate increase in area due to the edges compared to the floor is less than  $\approx 10\%$ . It should be noted that the phase noise error related to the edges is evenly distributed over the frequency domain, leading to a fairly insignificant increase in the overall distortion. In scenarios where the contribution from the edges is considered too dominant, it could be suppressed by increasing the time window utilized by the estimator. This is, for example, possible if the channel impulse response is much shorter than the CP, see [23]. However, such modifications come with an increased complexity.

The simulated average MSE over the OFDM symbol as a function of the phase noise level is shown for different cases in Figure 4. A low SNR level is shown in Figure 4 a), both to emphasize that phase noise is less of a problem when SNR is low, and to show that the performance of the proposed estimator, (28), is unaffected by the noise power level. A high SNR level, shown in Figure 4 b), indicates that advanced phase estimation methods are mainly important when the SNR is high. That is, when the payload consists of symbols with a high order of modulation, the sensitivity to phase noise is high. To achieve useful curves, the simulated MSE has been expressed as a loss compared to *BMSE with channel error* in dB scale. The simulated MSE of *Rx only* and *Tx/Rx* seems to match the BMSE over a large dynamic range of the phase noise for both low and high SNR, and only deviates at very high phase noise levels. The case *Genie Tx/Rx*, *linear model* is, as expected, almost completely overlapping the BMSE for all phase noise levels. For comparison, the MSE using low complexity CPE estimation<sup>6</sup> is included, denoted *CPE*, where

<sup>6</sup>Approximating the phase noise as a scalar phase rotation is well known and well studied in literature.

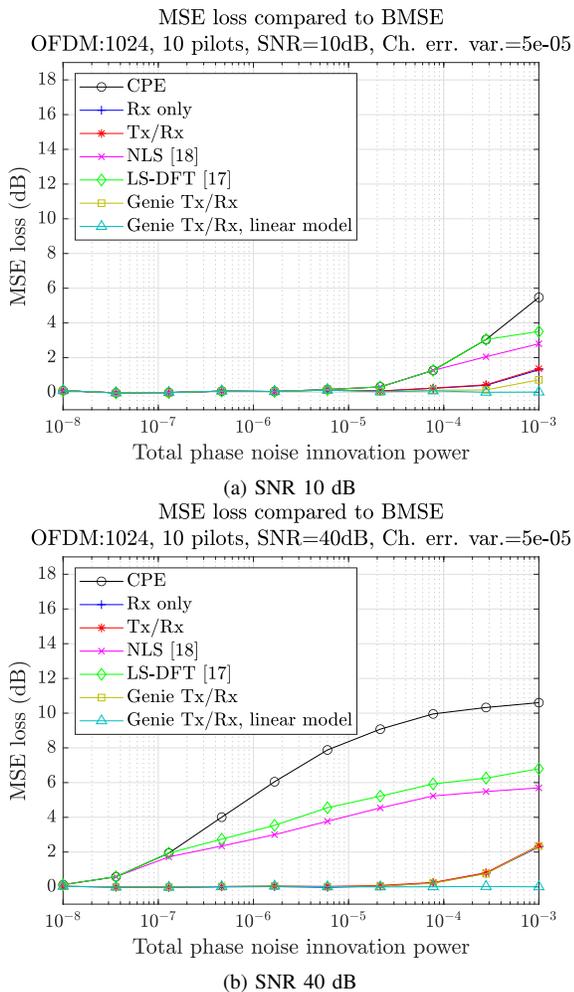


Fig. 4. Estimator MSE loss compared to *BMSE* with *channel error* for different cases at low and high SNR. The performance for *Rx only* and *Tx/Rx* is close to the *BMSE* over a wide dynamic range of the phase noise and SNR. Included, for comparison, are also a CPE estimator, the LS with DFT-base proposed by [17], and the NLS estimator proposed by [18].

the CPE is estimated with LS using the pilots. In addition, two other pilot-based methods from literature are also included, both of them derived for systems with phase noise only in the receiver (similarly to *Rx only*). First, the normalized LS (NLS) method proposed in [18], denoted *NLS*. As the phase noise is modeled using a number of constant segments in [18], the optimal number of segments is here found numerically for each phase noise level. Second, the LS method proposed by [17], denoted *LS-DFT*, which uses a subset of the DFT-basis to reduce complexity. As for *NLS*, the optimal number of parameters is found numerically for each phase noise level. Using a subset of the DFT-basis for approximating phase noise in OFDM applications is a common approach in the literature.

Simulated uncoded, phase noise degraded, BER as a function of SNR is compared with corresponding AWGN case, denoted *AWGN*, in Figure 5 for a configuration with 30 pilots and a phase noise innovation power of  $10^{-5}$ . The cases *Rx only* and *Tx/Rx* almost overlap with *Genie Tx/Rx*, which is reasonable when looking at Figure 4. At high SNR, *Rx only* has a slight offset compared to *Tx/Rx* and *Genie Tx/Rx*, and the

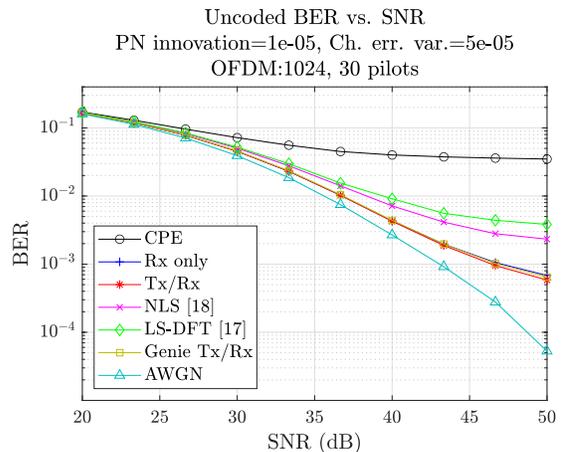


Fig. 5. Simulated uncoded BER as a function of SNR with a configuration with 30 pilots and a phase noise innovation power of  $10^{-5}$ . The cases *Rx only*, *Tx/Rx* and *Genie Tx/Rx* are almost completely overlapping, and the degradation compared to *AWGN* is small at low to medium SNR. Included, for comparison, are also a CPE estimator, the LS with DFT-base proposed by [17], and the NLS estimator proposed by [18].

reason for this offset is presented in conjunction with (23). For comparison, the three cases from the literature, defined above, are also shown. As discussed above, the performance of *NLS* and *LS-DFT* depends on the number of parameters used in the phase noise model, and the optimal number of parameters varies with the SNR and the phase noise level (here found numerically for each SNR). This is not a problem for the proposed MMSE estimator, (28). The BER degradation due to phase noise depends on the number of pilots, and an increased number of pilots will improve performance, i.e., "move" the BER curves closer to *AWGN*. However, an increased number of pilots also decreases the rate, and the optimal number of pilots to maximize the rate is discussed and analyzed below.

System performance can be analyzed in many different ways. Here, system performance is defined as the average rate of the payload subcarriers of the phase compensated received signal defined by (1), i.e., not after inverting the channel. The rate for payload subcarrier  $k$  is formulated as  $r_k = r_{\text{norm}} \log_2(1 + \text{SNDR}_k)$ , where  $r_{\text{norm}}$  is the payload ratio,  $r_{\text{norm}} = (N - N_{\text{pilot}})/N = N_{\text{payload}}/N$ , and  $\text{SNDR}_k$  is the SNDR for payload subcarrier  $k$ . The average rate for all payload subcarriers is  $1/N_{\text{payload}} \sum r_k$ . The subcarrier SNDR is found by simulation, where the noise corresponds to the total remaining error after phase noise compensation. The received signal power per subcarrier corresponds to the diagonal of  $\sigma_x^2 \mathbf{H} \mathbf{H}^H$  ( $\sigma_x^2$  is the transmitted power per subcarrier).

The rate will, of course, vary with the number of pilots. An increasing number of pilots will decrease the distortion related to the phase noise. However, as the pilot overhead increases, the rate decreases. For each phase noise level, the rate can be maximized with respect to the pilot-rate (number of pilots). The optimal relative rate (compared to the corresponding phase noise-free system), using *Rx only* estimation, as a function of phase noise is presented in Figure 6, where the maximum has been found using numerical search. Included is also the corresponding pilot-rate that maximizes

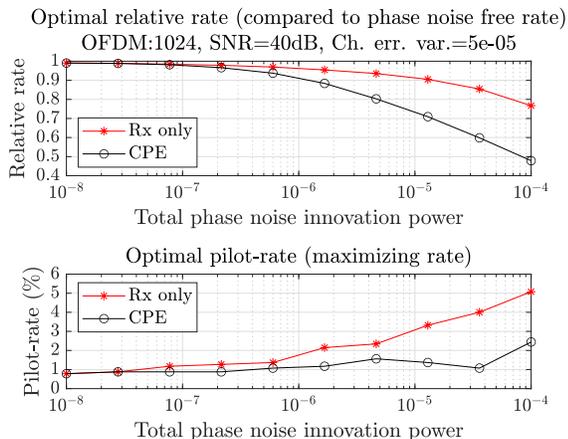


Fig. 6. Maximum relative rate (compared to the corresponding phase noise-free system) for *Rx only* and *CPE*, and the proposed method outperforms the *CPE* estimator over a large dynamic range. Included is also the corresponding pilot-rate (pilot overhead) that maximizes the rate for the two cases.

the rate. For comparison, the relative rate for a system with only *CPE* compensation is included, denoted *CPE*. As can be seen, *Rx only* outperforms *CPE* over a large dynamic range of phase noise. For an additional improved performance (higher rate), other processing is necessary, utilizing, for example, payload decisions. Since (17) allows for distributed pilots, any number of decisions is possible anywhere within the OFDM symbol, which differs from most decision directed proposals in literature.

It is interesting to compare the region with low to moderate rate loss (phase noise innovation power  $< 10^{-4}$ ) in Figure 6 with the corresponding performance of the estimator in Figure 4. Clearly, the approximations related to the higher order Taylor expansion terms, (13), when formulating the estimator have minor effect for realistic phase noise levels.

The complexity order of the LS and NLS, proposed by [17] and [18], is approximately  $N_{\text{pilot}}^3 + NN_{\text{pilots}}$ . The constrained LS algorithm proposed by [18] (not simulated in this paper), which showed better performance compared to NLS, has a much higher complexity order of  $N_{\text{pilot}}^{4.5} + NN_{\text{pilots}}$ . The proposed MMSE algorithm, (28), has a complexity order of approximately  $N_{\text{pilot}}^3 + NN_{\text{pilots}}$ . Looking at Figure 4, the performance of the proposed MMSE method is close to optimal for realistic phase noise levels over a large dynamic range of SNR, and outperforms the other methods with similar complexity.

## VII. CONCLUSION

In this paper, a novel pilot-based OFDM phase estimator has been demonstrated, where the pilots can be uniformly distributed in the frequency domain. The importance of utilizing the known statistics of the phase noise has been shown, and as can be seen in Figure 4, the estimator approaches a BMSE over a wide dynamic range of the phase noise level.

Although systems normally have phase noise in both the transmitter and the receiver, analysis shows that an estimator can assume that all phase noise completely allocated to the

receiver without significant degradation if the channel and phase noise meet certain conditions.

The rate of the system at a certain phase noise level can be maximized by using the optimal pilot-rate, as indicated in Figure 6. Hence, there exists a maximum rate for a certain phase noise level when using pilot-based phase noise estimation. For further increasing the rate, other techniques can be used in combination with the proposed pilot-based estimator, like, for example, payload decisions (where decisions are used together with pilot symbols). The proposed method sets no limits for which and how many payload symbols to utilize in the estimate.

Channels with Doppler have not been covered in this paper, and the proposed estimator is judged equally sensitive to Doppler, compared to existing methods for estimating phase noise in the literature. If Doppler is a significant part of the channel, it must be compensated before the proposed estimator. However, the proposed methodology might enable new possibilities to jointly handle several additional problems related to the channel, and these possibilities should be studied in the future. Additional research could, for example, include joint estimation of phase noise, and channel; joint estimation of phase noise, nonlinear distortion, and channel; and joint estimation of phase noise, nonlinear distortion, channel, and Doppler.

## APPENDIX A

### CPE NORMALIZED PHASE NOISE COVARIANCE

In this appendix, an analytical derivation of the covariance matrices of the zero mean, or *CPE*-normalized, phase noise,  $\mathbf{C}_{\theta}^{\text{Tx}}$  and  $\mathbf{C}_{\theta}^{\text{Rx}}$ , is presented.

The phase noise covariance of the *CPE*-normalized observation, (7), can be derived using the phase noise for sample  $l$ , formulated as

$$\phi_l = \alpha + \sum_{k=1}^l \Delta_k. \quad (36)$$

where  $\Delta_k$  is the phase noise innovation according to (6). The model describes a Wiener process of length  $l$  with an offset of  $\alpha$  (some unknown phase offset). The *CPE* is

$$\bar{\phi} = \frac{1}{N} \sum_{k=1}^N \phi_k. \quad (37)$$

The *CPE*-normalized phase  $l$  is  $\theta_l = \phi_l - \bar{\phi}$ , and it is straightforward to see that  $\mathbb{E}[\phi_l - \bar{\phi}] = 0$ . The *CPE*-normalized phase noise covariance matrix for both the transmitter and receiver has elements described by

$$\begin{aligned} \mathbf{C}_{\theta,k,l} &= \mathbb{E}[\theta_k \theta_l] = \mathbb{E}[\phi_k \phi_l] - \mathbb{E}[\phi_k \bar{\phi}] - \mathbb{E}[\phi_l \bar{\phi}] + \mathbb{E}[\bar{\phi} \bar{\phi}] \\ &= \mathbb{E}[\phi_k \phi_l] - \frac{1}{N} \mathbb{E} \left[ \phi_k \sum_{i=1}^N \phi_i \right] - \frac{1}{N} \mathbb{E} \left[ \phi_l \sum_{i=1}^N \phi_i \right] \\ &\quad + \frac{1}{N^2} \mathbb{E} \left[ \sum_{i=1}^N \phi_i \sum_{j=1}^N \phi_j \right]. \end{aligned} \quad (38)$$

The first term can be calculated using (6). The second and third term of (38), become, leaving out  $1/N$ ,

$$\begin{aligned} \mathbb{E} \left[ \phi_l \sum_{i=1}^N \phi_i \right] &= \mathbb{E} \left[ \sum_{j=1}^l \Delta_j \sum_{i=1}^N \sum_{m=1}^i \Delta_m \right] \\ &= \frac{l(l-1)}{2} \sigma_{\Delta}^2 + l(N-l+1) \sigma_{\Delta}^2 \\ &= \sigma_{\Delta}^2 \frac{l(2N+1) - l^2}{2} \end{aligned} \quad (39)$$

and the last term of (38), leaving out  $1/N^2$ , becomes

$$\begin{aligned} \mathbb{E} \left[ \sum_{i=1}^N \sum_{j=1}^N \phi_i \phi_j \right] &= \phi_1 \sum_{j=1}^N \phi_j + \phi_2 \sum_{j=1}^N \phi_j + \phi_3 \sum_{j=1}^N \phi_j + \dots \\ &= \sigma_{\Delta}^2 + \sigma_{\Delta}^2 + 2\sigma_{\Delta}^2(N-1) \\ &\quad + \sigma_{\Delta}^2 + 2\sigma_{\Delta}^2 + 3\sigma_{\Delta}^2(N-2) + \dots \\ &= \sigma_{\Delta}^2 \sum_{k=1}^N \sum_{i=1}^{k-1} i + \sigma_{\Delta}^2 \sum_{k=1}^N k(N-k+1) \\ &= \frac{\sigma_{\Delta}^2 N(2N^2 + 3N + 1)}{6}. \end{aligned} \quad (40)$$

The elements of the covariance matrix become

$$\begin{aligned} C_{\theta,k,l} &= \max(k,l) \sigma_{\Delta}^2 + \sigma_{\Delta}^2 \frac{l^2 - l(2N+1)}{2N} \\ &\quad + \sigma_{\Delta}^2 \frac{k^2 - k(2N+1)}{2N} + \sigma_{\Delta}^2 \frac{2N^2 + 3N + 1}{6N}. \end{aligned} \quad (41)$$

#### APPENDIX B

##### REDUCTION OF PHASE NOISE MODEL COMPLEXITY

In this appendix, a method for reducing the complexity of phase noise estimators is presented, where the approximation allows a representation of the phase noise over an OFDM symbol using significantly fewer parameters than the number of OFDM subcarriers,  $N$ . The method is similar to what is normally done using PCA.

The covariance matrix of the CPE-normalized transmitter or receiver phase noise can be factorized as

$$\mathbf{C}_{\theta} = \mathbf{C}_{\tilde{\theta}}^{\frac{1}{2}} \mathbf{C}_{\tilde{\theta}}^{\frac{1}{2}\text{T}}, \quad (42)$$

where the factorization is performed using the singular value decomposition (SVD) into  $\mathbf{C}_{\theta} = \mathbf{U}_{\theta} \mathbf{\Sigma}_{\theta} \mathbf{V}_{\theta}^{\text{H}}$ . It should be noted that  $\mathbf{U}_{\theta} = \mathbf{V}_{\theta}$  since  $\mathbf{C}_{\theta}$  is symmetric. The square root matrix is identified as  $\mathbf{C}_{\tilde{\theta}}^{\frac{1}{2}} = \mathbf{U}_{\theta} \mathbf{\Sigma}_{\theta}^{\frac{1}{2}}$ . Phase noise is real-valued, and the left and right singular matrices are therefore real-valued. A model of the phase noise can be formulated as

$$\boldsymbol{\theta} = \mathbf{U}_{\theta} \mathbf{\Sigma}_{\theta}^{\frac{1}{2}} \mathbf{v}, \quad (43)$$

where  $\mathbf{v}$  is  $\mathcal{N}(\mathbf{0}, \mathbf{I})$ . It is straightforward to see that  $\mathbb{E}[\mathbf{U}_{\theta} \mathbf{\Sigma}_{\theta}^{\frac{1}{2}} \mathbf{v} (\mathbf{U}_{\theta} \mathbf{\Sigma}_{\theta}^{\frac{1}{2}} \mathbf{v})^{\text{T}}] = \mathbf{C}_{\theta}$ . Hence, the correlated phase noise can be modeled as a linear combination of the left orthonormal singular vectors of the phase noise covariance matrix  $\mathbf{C}_{\theta}$ . The contribution of the different singular vectors

is scaled by the elements of the diagonal matrix  $\mathbf{\Sigma}_{\theta}^{\frac{1}{2}}$ . The singular values of  $\mathbf{\Sigma}_{\theta}$  are in this case far from equal and the strongest will contribute the most.

The phase noise can be approximated by using the singular vectors corresponding to the strongest singular values as

$$\boldsymbol{\theta} \approx \mathbf{U}_{\tilde{\theta}}^k \tilde{\boldsymbol{\theta}}_k, \quad (44)$$

where the  $N \times k$  matrix  $\mathbf{U}_{\tilde{\theta}}^k$  contains the  $k$  singular vectors corresponding to the  $k$  strongest singular values and  $\tilde{\boldsymbol{\theta}}_k$  contains  $k$  phase noise parameters. The covariance matrix of  $\tilde{\boldsymbol{\theta}}_k$  is formulated as

$$\mathbf{C}_{\tilde{\theta}} = \mathbf{U}_{\tilde{\theta}}^{\text{T}} \mathbf{C}_{\theta} \mathbf{U}_{\tilde{\theta}}^k, \quad (45)$$

which is a  $k \times k$  diagonal matrix. Hence, using PCA, the correlated phase noise has been transformed into a smaller set of uncorrelated phase noise parameters, where the variance of the new set of parameters is maximized.

#### APPENDIX C

##### PARAMETER AND ERROR COVARIANCE

In this appendix, the parameter covariance,  $\mathbf{C}_{\theta_{\text{tot}}}$ , in (28) is derived. Furthermore, the covariance of the different parts of the combined noise,  $\mathbf{C}_{w_{\text{tot}}}$ , in (28) is formulated. In particular, a simplification of the covariance of  $\mathbf{r} + \mathbf{n}$  is presented and an empirical alternative is proposed, containing only a few system-dependent parameters. The analysis assumes that the contribution of the error term  $\mathbf{e}$ , (3), is close to zero.

The CPE phase,  $\bar{\theta}$ , is uniform  $[0 \ 2\pi]$  and uncorrelated with the zero average phase noise, and the covariance of the vector variable  $[\alpha \ \beta]^{\text{T}}$  becomes<sup>7</sup>  $1/2 \mathbf{I}_2$ . Using (8) together with the covariance of  $\alpha$  and  $\beta$ , the covariance matrix of  $\boldsymbol{\theta}_{\text{tot}}$  in (24) becomes

$$\mathbf{C}_{\theta_{\text{tot}}} = \begin{pmatrix} \frac{1}{2} \mathbf{I}_2 & \mathbf{0}_{2 \times N} & \mathbf{0}_{2 \times N} \\ \mathbf{0}_{N \times 2} & \mathbf{C}_{\theta}^{\text{Tx}} + \mathbf{C}_{\theta}^{\text{Rx}} & \mathbf{C}_{\theta}^{\text{Tx}} \\ \mathbf{0}_{N \times 2} & \mathbf{C}_{\theta}^{\text{Tx}} & \mathbf{C}_{\theta}^{\text{Tx}} \end{pmatrix}, \quad (46)$$

since the transmitter and receiver phase noise are independent and the CPE parameters is uncorrelated with the phase noise. The matrix  $\mathbf{0}_{k \times l}$  is an all zero matrix with  $l$  columns and  $k$  rows, and  $\mathbf{I}_k$  is an identity matrix of size  $k \times k$ , and  $\mathbf{C}_{\theta}^{\text{Tx}}$  respectively  $\mathbf{C}_{\theta}^{\text{Rx}}$  are the  $N \times N$  covariance matrices of the CPE-normalized transmitter and receiver phase noise. Hence, the size of  $\mathbf{C}_{\theta_{\text{tot}}}$  is  $(2N+2) \times (2N+2)$ .

The combined channel estimation error and the AWGN,  $\Delta_H \mathbf{x} + \mathbf{w}$ , in (24) has, using, (5), a covariance matrix formulated as

$$\begin{aligned} \mathbf{C}_{\Delta_H \mathbf{x} + \mathbf{w}} &= \mathbb{E} \left[ \Delta_H \underbrace{\left( e^{i\bar{\theta}} \mathbf{x} + \mathbf{F} \mathbf{R}_{\hat{H}^{-1}y} \boldsymbol{\theta}_{\text{Tx}} \right)}_{\approx e^{i\bar{\theta}} \mathbf{x}} (\dots)^{\text{H}} \Delta_H^{\text{H}} \right] \\ &\quad + \sigma_w^2 \mathbf{I} \\ &\approx \sigma_x^2 \sigma_{\Delta}^2 \mathbf{I} + \sigma_w^2 \mathbf{I}, \end{aligned} \quad (47)$$

<sup>7</sup>The diagonal of the covariance matrix is formed from the integral over  $[0 \ 2\pi]$  of  $\cos^2(\bar{\theta})/(2\pi) = (1 + \cos(2\bar{\theta}))/(4\pi)$  and  $\sin^2(\bar{\theta})/(2\pi) = (1 - \cos(2\bar{\theta}))/(4\pi)$ , which is  $1/2$ . The non-diagonal elements are the integral over  $[0 \ 2\pi]$  of  $\cos(\bar{\theta}) \sin(\bar{\theta})/(2\pi) = (\sin(0) + \sin(2\bar{\theta}))/(4\pi)$ , which is 0.

since they are uncorrelated. The phase noise part is assumed to be much smaller than the signal.

The error term  $\mathbf{r} + \mathbf{n}$  in (24) is complicated to analyze and it is difficult to formulate an analytical error covariance. However, by simplifying the expression, as shown in Section D, it is possible to formulate an empirical expression, modeling the covariance matrix using only the channel and the innovation variance of the transmitter and receiver phase noise. Assuming the phase noise small, the approximate covariance can, using (55), be expressed by evaluating

$$\begin{aligned} \mathbf{r} + \mathbf{n} \approx & -i\mathbf{F}\frac{(i\Theta_{\text{Rx}})^2}{2}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + \mathbf{H}\mathbf{F}\frac{(i\Theta_{\text{Tx}})^2}{2}\mathbf{F}^{-1}\mathbf{x} \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{H}\mathbf{x} \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{w}. \end{aligned} \quad (48)$$

Five different contributions can be identified, and the channel,  $\mathbf{H}$ , shapes the first two.

The data,  $\mathbf{x}$ , and the AWGN,  $\mathbf{w}$ , have diagonal covariance matrices, and there are reasons to believe that covariance matrix of (48) is diagonal, where one part is shaped by the channel<sup>8</sup> and one part is white (from the last three terms of (48)). The empirical covariance<sup>9</sup> is postulated to be

$$\begin{aligned} \mathbf{C}_{\mathbf{r}+\mathbf{n}} \approx & \xi\mathbf{H}\mathbf{H}^{\text{H}}(\sigma_{\Delta_{\text{Tx}}}^4 + \sigma_{\Delta_{\text{Rx}}}^4) \\ & + \zeta\sigma_{\Delta_{\text{Tx}}}^2\sigma_{\Delta_{\text{H}}}^2\mathbf{I} + \gamma\sigma_{\Delta_{\text{Rx}}}^2\sigma_w^2\mathbf{I}, \end{aligned} \quad (49)$$

where three unknown scaling parameters,  $\xi$ ,  $\zeta$  and  $\gamma$ , have been introduced. The values of these parameters can be found using the simulated covariance of  $\mathbf{r} + \mathbf{n}$ , where  $\mathbf{n}$  can be generated from (15) by subtracting the linear terms without AWGN, and  $\mathbf{r}$  can be generated from (16) using (18). The empirical covariance, (49), and the simulated covariance based on  $\mathbf{r} + \mathbf{n}$  are compared in Section VI.

#### APPENDIX D SIMPLIFICATIONS OF APPROXIMATION ERRORS

This appendix contains simplifications of  $\mathbf{r} + \mathbf{n}$ , (13) and (16). The simplifications are performed in two steps, first by separately approximating  $\mathbf{r}$ , followed by simplifying the sum of the approximated  $\mathbf{r}$  and  $\mathbf{n}$ . The CPE phasor has been left out because it will not affect the covariance. The cyclic prefix term  $\mathbf{e}$ , (3), has also been left out since it is normally significantly smaller than  $\mathbf{w}$ ,  $\mathbf{r}$  and  $\mathbf{n}$ .

To simplify the analysis of  $\mathbf{r}$ , the notation of the time domain diagonal matrix,  $\mathbf{R}$ , is replaced by the notation used

<sup>8</sup>The modeling can be understood by, for example, looking at the covariance contribution of the first term is  $1/4\mathbb{E}[\mathbf{F}\Theta_{\text{Rx}}^2\mathbf{F}^{-1}\mathbf{H}\mathbf{x}\mathbf{x}^{\text{H}}\mathbf{H}^{\text{H}}\mathbf{F}^{-\text{H}}\Theta_{\text{Rx}}^2\mathbf{F}^{\text{H}}]$ . The left and right multiplication of the diagonal phase noise matrices should result in a diagonal matrix when taking the mean since the data symbols,  $\mathbf{x}$ , are independent.

<sup>9</sup>Here the square cross term have been left out since it is small in comparison.

in (12). Using (16) with (4) and (18), the resulting expression becomes

$$\begin{aligned} \mathbf{r} = & i(\mathbf{F}\mathbf{R}_{\text{Hx}}\mathbf{H}\mathbf{F}\mathbf{R}_{\text{x}})\begin{pmatrix} \theta_{\text{Rx}} \\ \theta_{\text{Tx}} \end{pmatrix} \\ & - i(\mathbf{F}\mathbf{R}_{\text{y}}\mathbf{H}\mathbf{F}\mathbf{R}_{\hat{\text{H}}^{-1}\text{y}})\begin{pmatrix} \theta_{\text{Rx}} \\ \theta_{\text{Tx}} \end{pmatrix} - i\Delta_{\text{H}}\mathbf{F}\mathbf{R}_{\hat{\text{H}}^{-1}\text{y}}\theta_{\text{Tx}} \\ = & i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{y}) + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}(\mathbf{x} - \hat{\mathbf{H}}^{-1}\mathbf{y}) \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\hat{\mathbf{H}}^{-1}\mathbf{y}. \end{aligned} \quad (50)$$

Replacing  $\mathbf{y}$  with (12) results in

$$\begin{aligned} \mathbf{r} = & i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}(\mathbf{H}\mathbf{x} - \mathbf{H}\mathbf{x} \\ & - i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} - i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - \mathbf{n} - \mathbf{w}) \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}(\mathbf{x} - \hat{\mathbf{H}}^{-1} \\ & \times \underbrace{(\mathbf{H}\mathbf{x} + i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} + \mathbf{n} + \mathbf{w})}_{\approx \mathbf{H}\mathbf{x}}) \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\hat{\mathbf{H}}^{-1} \\ & \times \underbrace{(\mathbf{H}\mathbf{x} + i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} + \mathbf{n} + \mathbf{w})}_{\approx \mathbf{H}\mathbf{x}}, \end{aligned} \quad (51)$$

where the approximations have been highlighted.

The inverse of the channel estimate, (4), can be approximated if the estimation error is small, resulting in

$$\begin{aligned} \hat{\mathbf{H}}^{-1} = & (\mathbf{I} + \mathbf{H}^{-1}\Delta_{\text{H}})^{-1}\mathbf{H}^{-1} \\ \approx & (\mathbf{I} - \mathbf{H}^{-1}\Delta_{\text{H}})\mathbf{H}^{-1} \\ = & \mathbf{H}^{-1} - \mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{H}^{-1}, \end{aligned} \quad (52)$$

using  $(\mathbf{I} + \mathbf{A})^{-1} \approx (\mathbf{I} - \mathbf{A})$  if the singular values of  $\mathbf{A}$  are  $\ll 1$ . Using the approximated channel inverse,  $\mathbf{r}$  becomes

$$\begin{aligned} \mathbf{r} \approx & i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1} \\ & \times \left( \underbrace{-i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} - i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - \mathbf{n} - \mathbf{w}}_{\approx -\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} - i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - \mathbf{w}} \right) \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1} \left( \mathbf{x} - \underbrace{(\mathbf{H}^{-1} - \mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{H}^{-1})\mathbf{H}\mathbf{x}}_{=\mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{x}} \right) \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1} \left( \mathbf{H}^{-1} - \mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{H}^{-1} \right) \mathbf{H}\mathbf{x} \\ & \quad \quad \quad \approx i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} \\ \approx & i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1} \left( -i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} - i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - \mathbf{w} \right) \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{H}\mathbf{x} \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x}, \end{aligned} \quad (53)$$

where simplifications and approximations have been highlighted.

The evaluation of  $r$  by itself ends at this point. The combination of the errors,  $r + n$ , can thus be approximately formulated as

$$\begin{aligned} r + n \approx & -i\mathbf{F}(i\Theta_{\text{Rx}})^2\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + \mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{x} \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{w} \\ & + \mathbf{F}\frac{(i\Theta_{\text{Rx}})^2}{2}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + \mathbf{H}\mathbf{F}\frac{(i\Theta_{\text{Tx}})^2}{2}\mathbf{F}^{-1}\mathbf{x} \\ & - \mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x}, \end{aligned} \quad (54)$$

approximating (13) with the initial second order terms. The resulting approximation becomes

$$\begin{aligned} r + n \approx & -i\mathbf{F}\frac{(i\Theta_{\text{Rx}})^2}{2}\mathbf{F}^{-1}\mathbf{H}\mathbf{x} + \mathbf{H}\mathbf{F}\frac{(i\Theta_{\text{Tx}})^2}{2}\mathbf{F}^{-1}\mathbf{x} \\ & + i\mathbf{H}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{H}^{-1}\Delta_{\text{H}}\mathbf{x} \\ & - i\Delta_{\text{H}}\mathbf{F}\Theta_{\text{Tx}}\mathbf{F}^{-1}\mathbf{x} - i\mathbf{F}\Theta_{\text{Rx}}\mathbf{F}^{-1}\mathbf{w}. \end{aligned} \quad (55)$$

## REFERENCES

- [1] U. Gustavsson, P. Frenger, C. Fager, T. Eriksson, H. Zirath, F. Dielacher, C. Studer, A. Parssinen, R. Correia, J. N. Matos, D. Belo, and N. B. Carvalho, "Implementation Challenges and Opportunities in Beyond-5G and 6G Communication," *IEEE Journal of Microwaves*, vol. 1, pp. 86–100, 2021.
- [2] T. H. Lee and A. Hajimiri, "Oscillator phase noise: A tutorial," *IEEE Journal of Solid-State Circuits*, vol. 35, pp. 326–335, 2000.
- [3] A. Lapidoth, *A Foundation In Digital Communication*, 2nd ed. Cambridge University Press, 2017.
- [4] J. P. Santacruz, S. Rommel, U. Johannsen, A. Jurado-Navas, and I. T. Monroy, "Analysis and Compensation of Phase Noise in Mm-Wave OFDM ARoF Systems for beyond 5G," *Journal of Lightwave Technology*, vol. 39, pp. 1602–1610, 3 2021.
- [5] S. T. Le, T. Kanesan, E. Giacomidis, N. J. Doran, and A. D. Ellis, "Quasi-pilot aided phase noise estimation for coherent optical OFDM systems," *IEEE Photonics Technology Letters*, vol. 26, pp. 504–507, 3 2014.
- [6] S. T. Le, P. A. Haigh, A. D. Ellis, and S. K. Turitsyn, "Blind Phase Noise Estimation for CO-OFDM Transmissions," *Journal of Lightwave Technology*, vol. 34, pp. 745–753, 1 2016.
- [7] S. Wu and Y. Bar-Ness, "A phase noise suppression algorithm for OFDM-based WLANs," *IEEE Communications Letters*, vol. 6, pp. 535–537, 12 2002.
- [8] F. Munier, T. Eriksson, and A. Svensson, "Receiver algorithms for OFDM systems in phase noise and AWGN," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*, vol. 3, pp. 1998–2002, 2004.
- [9] W. Rave, D. Petrovic, and G. Fettweis, "Iterative Correction of Phase Noise in Multicarrier Modulation," *Proc. 9th Int. OFDM Workshop*, 2004.
- [10] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: characterization and compensation," *IEEE Transactions On Communications*, vol. 55, 2007.
- [11] S. Wu, P. Liu, and Y. Bar-Ness, "Phase noise estimation and mitigation for OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 3616–3625, 12 2006.
- [12] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Phase Noise Estimation in OFDM: Utilizing Its Associated Spectral Geometry," *IEEE Transactions on Signal Processing*, vol. 64, pp. 1999–2012, 4 2016.
- [13] S. Bittner, W. Rave, and G. Fettweis, "Joint Iterative Transmitter and Receiver Phase Noise Correction using Soft Information," *International Conference on Communications*, 2007.
- [14] G. Liu and W. Zhu, "Compensation of phase noise in OFDM systems using an ICI reduction scheme," *IEEE Transactions on Broadcasting*, vol. 50, pp. 399–407, 12 2004.
- [15] V. Syrjala, T. Levanen, T. Ihalainen, and M. Valkama, "Pilot Allocation and Computationally Efficient Non-Iterative Estimation of Phase Noise in OFDM," *IEEE Wireless Communications Letters*, vol. 8, pp. 640–643, 4 2019.
- [16] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "A non-iterative technique for phase noise ICI mitigation in packet-based OFDM systems," *IEEE Transactions on Signal Processing*, vol. 58, pp. 5945–5950, 11 2010.
- [17] R. A. Casas, S. L. Biracree, and A. E. Youtz, "Time domain phase noise correction for OFDM signals," *IEEE Transactions on Broadcasting*, vol. 48, pp. 230–236, 2002.
- [18] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Constrained phase noise estimation in OFDM using scattered pilots without decision feedback," *IEEE Transactions on Signal Processing*, vol. 65, pp. 2348–2362, 5 2017.
- [19] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: a unifying theory and numerical methods for characterization," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, pp. 655–674, 2000.
- [20] S. Mandelli, M. Magarini, and A. Spalvieri, "Modeling the filtered and sampled continuous-time signal affected by wiener phase noise," *2014 19th European Conference on Networks and Optical Communications, NOC 2014*, pp. 173–178, 12 2014.
- [21] H. Abdi and L. J. Williams, "Principal component analysis," pp. 433–459, 7 2010.
- [22] S. M. Kay, *Statistical Signal Processing: Estimation Theory*, 20th ed. Prentice Hall PTR, 2013.
- [23] C. Y. Ma, C. Y. Wu, and C. C. Huang, "A simple ICI suppression method utilizing cyclic prefix for OFDM systems in the presence of phase noise," *IEEE Transactions on Communications*, vol. 61, pp. 4539–4550, 11 2013.



**Björn Gävert** received the M.S. degree in Electrical Engineering from Chalmers University of Technology, Gothenburg, Sweden, in 1996. After that, he started working for Ericsson AB and he has had several different work positions, such as software designer, hardware designer, and radio system engineer. He was appointed Radio Signal Processing Senior Specialist in 2010, and became responsible for the digital signal processing architecture in the Ericsson MINI-LINK product family. Within the senior specialist role, he performed research on several

different topics, for example, power amplifier digital pre-distortion, digital post-compensation, and carrier phase estimation. From 2018 he became part of a team researching radio base station up-link digital hardware compensation (homodyne receivers, ADC, etc.).

He is currently an ind. Ph.D. student working towards the Ph.D. degree in electrical engineering.



**Mikael Coldrey** received the M.Sc. degree in applied physics and electrical engineering from Linköping University, Linköping, Sweden, in 2000, and the Ph.D. degree in electrical engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2006. He joined Ericsson Research in 2006, where he is currently a Principal Researcher. Between 2017-2020 he was Head of the microwave systems research section at Ericsson Research. Mikael received the Docent (habilitation) promotion from Chalmers in 2020, and was the same

year appointed Adjunct Professor in Communication Systems at Chalmers. His main research interests are in the areas of wireless communications, signal processing, advanced antenna systems, and millimeter wave communications for radio access and wireless transport systems.



**Thomas Eriksson** received the Ph.D. degree in Information Theory from Chalmers University of Technology, Gothenburg, Sweden, in 1996. From 1990 to 1996, he was at Chalmers. In 1997 and 1998, he was at AT&T Labs - Research, Murray Hill, NJ, USA. In 1998 and 1999, he was at Ericsson Radio Systems AB, Kista, Sweden. Since 1999, he has been with Chalmers University, where he is currently a professor of communication systems. Further, he was a guest professor with Yonsei University, S. Korea, in 2003-2004. He has authored or co-authored

more than 300 journal and conference papers, and holds 14 patents.

Prof. Eriksson is leading the research on hardware-constrained communications with Chalmers University of Technology. His research interests include communication, data compression, and modeling and compensation of non-ideal hardware components (e.g. amplifiers, oscillators, and modulators in communication transmitters and receivers, including massive MIMO). Currently, he is leading several projects on e.g. 1) massive MIMO communications with imperfect hardware, 2) wideband linearization (DPD), 3) Efficient and linear transceivers, etc. He is currently the Vice Head of the Department of Signals and Systems with Chalmers University of Technology, where he is responsible for undergraduate and master's education.