IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received February 28, 2020, accepted March 4, 2020, date of publication March 9, 2020, date of current version March 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2979376

Phase Noise Compensation for CFBMC–OQAM Systems Under Imperfect Channel Estimation

LONG D. LE AND HA H. NGUYEN, (Senior Member, IEEE)
Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada

Corresponding author: Long D. Le (long.le@usask.ca)

This work was supported by the University of Saskatchewan's Dean's Scholarship and the Natural Sciences and Engineering Research Council (NSERC) Discovery Grant.

ABSTRACT Among many multi-carrier systems, circular filter-bank multi-carrier offset quadrature amplitude modulation (CFBMC-OQAM) is one of promising candidates for future wireless communications. This paper studies the impact of phase noise and its compensation for CFBMC-OQAM under imperfect channel estimation, which has not been done before. In the presence of phase noise, a two-stage phase noise compensation algorithm is proposed. In the first stage, the channel frequency response and phase noise are estimated based on the transmission of a preamble. Such a preamble is designed to minimize the channel mean squared error. In the second stage, the estimated channel obtained from the first stage together with pilot symbols are used to compensate for the phase noise and detect the transmitted signal. Simulation results obtained under practical scenarios show that the proposed algorithm effectively estimates the channel frequency response and compensates for the phase noise. The proposed algorithm is also shown to outperform an existing algorithm that performs iterative phase noise compensation when phase noise impact is high.

INDEX TERMS FBMC, OFDM, CFBMC-OQAM, phase noise, preamble design, channel estimation.

I. INTRODUCTION

Multi-carrier transmission techniques have been extensively studied over the last few decades [1]. Currently, cyclic prefix orthogonal frequency-division multiplexing (CP-OFDM)[1] [1], [2] is a popular choice for broadband wireless communication systems because it offers many advantages including resistance to multipath distortions by using CP, and a simple one-tap channel equalization, and efficient implementations based on the fast Fourier transform (FFT). However, OFDM also has its own disadvantages. Two major disadvantages are: (i) the effect of rectangular pulse shaping in the conventional OFDM leads to a high out-of-band emission to neighboring frequency bands [1], and (ii) fairly complicated carrier frequency offset and timing synchronization techniques are needed to establish subcarrier orthogonality at the receiver [3]–[7].

In the search for more favorable waveforms for the physical layer of the next-generation wireless networks, a number of signaling methods has been suggested. The proposed waveforms can be broadly classified into two groups, one with linear pulse shaping and the other with circular pulse shaping. Filter-bank multi-carrier offset quadrature amplitude modulation (FBMC-OQAM) is one of the signaling methods in the first group that can resolve the above problems in OFDM [1], [8]. Specifically, with the use of well-designed prototype filters [9]–[11], an FBMC system can avoid the high spectral leakage issue in the conventional OFDM. On the other hand, using OQAM splits complex data into real and imaginary parts which consequently relaxes the orthogonality condition to the real field [1], [8]. However, FBMC-OQAM does not have a block-based structure like OFDM, which makes the channel estimation and equalization tasks significantly more complicated when compared to OFDM.

Recently proposed as candidates for the air interface of 5G networks [12], generalized frequency division multiplexing (GFDM) [13], [14] and circular filter-bank multi-carrier offset quadrature amplitude modulation (CFBMC-OQAM) [15] are signaling methods that use circular pulse shaping. With circular pulse shaping, the length of the CP can be chosen to be the same as the length of the channel impulse response, i.e., independent of the length of the added pulse shaping filter. This is not the case with linear pulse shaping in which the CP length needs to cover both the length of the channel impulse response and the delay introduced by the

The associate editor coordinating the review of this manuscript and approving it for publication was Yunlong Cai.

[1]For brevity, CP-OFDM is simply referred to as OFDM in the paper.

shaping filter. Similar to the conventional OFDM, the channel estimation and equalization in these systems can be simply performed after removing the CP. However, different from GFDM which is a non-orthogonal system and thus can severely suffer from the impact of inter-symbol interference (ISI) and inter-carrier interference (ICI), CFBMC-OQAM is an orthogonal system in the real domain.

In practice, to perform the channel equalization task, the channel state information has to be first estimated based on the transmission of either pilots or preambles as in [16]–[18] for OFDM, and [19]–[22] for FBMC-OQAM, and [23], [24] for GFDM. However, to the best of our knowledge, all studies on CFBMC-OQAM assume that the channel frequency response is perfectly known for the equalization task [12], [15] and thus the bit-error rate (BER) performance of CFBMC-OQAM is equivalent to that of OFDM under the same system configuration. Such an assumption does not hold in practice.

Any multi-carrier systems suffer from physical impairments, especially the phase noise (PN). Impact of PN on OFDM systems has been extensively studied in the last few decades [25]–[29]. In the time domain, the effect of PN can be represented as a multiplicative noise in the form of a complex exponential function. In the frequency domain, on the other hand, the PN impact can be separated into two components, namely common phase error (CPE) and ICI. The first component results in a common rotation of all symbols, while the second component is different for every symbol. Mitigating the PN impact in OFDM systems can be performed in either the frequency domain [26], [27], [30]–[32] or the time domain [28], [29], [33]–[37].

In [31], the frequency-domain PN vector is estimated by solving a constrained quadratic optimization problem. However, the problem is solved under the assumption that PN is very small so that the constraint is relaxed and applied only for the real part of the first frequency-domain PN component. To improve the PN estimation in [31], better constraints are considered in [32] and [34]. The constraint in [32] is derived based on a geometrical structure associated with the frequency-domain PN components, whereas the constraint in [34] relies on the fact that the complex exponential functions of PN have unit magnitudes. Applying the so-called S-procedure, the optimization problem in [32] is solved as a semidefinite programming problem. However, because of the inherently high computational complexity of semidefinite programming, the algorithm in [32] may not work efficiently when the number of sub-carriers is very large as expected in future communication systems. On the other hand, an algorithm known as majorization-minimization [38] is applied to estimate the channel and PN in [34] with low complexity that allows to deal with systems having a much larger number of sub-carriers.

In [35], an iterative PN compensation algorithm is proposed for OFDM. With this iterative approach, based on the estimated channel and the output of a soft-decision Viterbi decoder, the receiver tries to replicate the PN-free transmitted signal and uses it to estimate the PN in the time domain by a one-tap least mean square (LMS) method. Performance advantage of the iterative PN compensation algorithm over other algorithms is illustrated in terms of the packet-error rate (PER) in [35]. The iterative technique in [35] is further considered in [36] and [37] for 60 GHz and 10 Gbps OFDM experimental systems, respectively.

The impact of PN has recently been considered in [39], [40] for FBMC-OQAM, [41] for GFDM, and [42] for CFBMC-OQAM. Different from OFDM, FBMC-OQAM, GFDM and CFBMC-OQAM suffer from two intrinsic interferences, namely ISI and ICI. Therefore, these systems generally have additional interference components when PN is taken into account. Consequently, mitigating the impact of PN in such systems is more complicated than that in OFDM [39]. In [40], two novel CPE compensation algorithms, one is a pilot-based approach and the other is a blind approach exploiting the interference structure from the neighboring sub-channels, are proposed for FBMC-OQAM. Under the small PN assumption, the paper shows that the proposed algorithms nearly achieve the optimal performance when no PN is considered. The authors in [41] propose a filter for GFDM systems to maximize the signal-to-interference ratio in the presence of timing offset, carrier frequency offset and PN. In our previous study [42], we investigated the impact of two different PN types, one is based on a free-running (FR) oscillator and the other is based on a phase-locked loop (PLL) oscillator, on CFBMC-OQAM systems and proposed an algorithm to effectively compensate the PN impact by taking the intrinsic self-interferences into account. However, the work in [42] assumes that the channel frequency response is perfectly known.

The impact of PN on CFBMC-OQAM and its compensation under imperfect channel estimation shall be studied in this paper. Since CFBMC-OQAM is different from both OFDM and FBMC in terms of signal processing operations, existing techniques developed for OFDM and FBMC might not provide good performance when applied straightforwardly to CFBMC-OQAM. First, it shall be shown that, compared to OFDM, performance of CFBMC-OQAM is much more susceptive to imperfect channel estimation even when there is no PN. The paper then proposes a preamble design to minimize the channel mean squared error (MSE) and consequently mitigate the performance degradation. In the presence of PN, the paper proposes a two-stage PN compensation algorithm. With the designed preamble, channel and PN are estimated in the first stage. Different from previous works [31], [32], [34] where the channel is estimated based on the least square (LS) method, the channel estimation in this paper is performed based on the minimum mean-squared error (MMSE) method to significantly improve the estimation performance. In the second stage, the estimated channel obtained from the first stage is then utilized together with pilots to compensate for the impact of PN and detect the transmitted data. To avoid error propagation and long latency experienced by iterative PN compensation

algorithms [35]–[37], a non-iterative method is proposed in the second stage. Simulation results obtained under practical scenarios illustrate the effectiveness of the proposed method in compensating for the PN impact in the presence of imperfect channel estimation. Especially, the proposed two-stage algorithm is shown to outperform the iterative PN compensation algorithm in [35] when the PN is large.

It is relevant to point out that paper [43] proposes a two-step algorithm to estimate the carrier frequency offset (CFO) and channel in a multi-input multi-output FBMC-OQAM system. Specifically, in the first step, the CFO is coarsely estimated by using the cross correlation between the received signal and the known preamble. After CFO compensation, the effective channel, which is a combination of the residual CFO and the real channel, is estimated in the second step by either weighted LS or MMSE estimator. The difference between [43] and our paper is that the impact of CFO is represented via a single unknown variable. Once an estimated CFO is obtained from the first step based on preambles, it is used for channel estimation in the second step and data detection for the remaining transmission blocks. However, in our paper, PN needs to be modeled as a random process and it changes from one transmission block to the other. As such, the estimated PN vector obtained based on a preamble block cannot be reused for the data detection in the next blocks.

This paper is organized as follows. The system model is presented in Section II, which includes a brief review of a CFBMC-OQAM system and characteristics of two PN types. In the first part of Section III, a preamble design is proposed to mitigate the performance degradation in CFBMC-OQAM due to imperfect channel estimation and when PN is absent. Under the presence of PN, the paper proposes a two-stage PN compensation algorithm. The first stage is presented in the second part of Section III to estimate the channel and PN. The second stage is described in Section IV to compensate for the PN impact and detect the transmitted data. Simulation results are provided in Section V. Section VI concludes the paper.

Notation: Lowercase letters are used to denote scalars. Lowercase boldface and uppercase boldface letters stand for vectors and matrices, respectively. Symbols $(\cdot)^{\dagger}$, $(\cdot)^{T}$, $(\cdot)^{H}$, $(\cdot)^{-1}$, $(\cdot)^{*}$, $|\cdot|$, $||\cdot||_2$, and $\circledast$ denote pseudo-inverse, transpose, Hermitian, inverse, complex conjugate, modulus, ℓ_2-norm and circular convolution operations, respectively. Symbols $\Re\{\cdot\}$, $\Im\{\cdot\}$, diag$[\cdot]$, and E$[\cdot]$ denote the real and imaginary parts of a complex variable, a diagonal matrix if the argument is a vector, or a vector if the argument is a diagonal matrix, and expectation operation, respectively. $\mathbf{1}_{N\times 1}$ denotes a column vector whose N components are all 1's.

II. SYSTEM MODEL

A. CFBMC-OQAM

Consider a discrete-time complex baseband equivalent CFBMC-OQAM system as illustrated in Fig. 1. Let $N = KM$. Then, the transmitted signal vector corresponding to the ith transmission block $\mathbf{x}_i = \left[x_i(0)\ x_i(1)\ \cdots\ x_i(N-1)\right]^T$ is constructed from a $K \times 2M$ block of data defined over K sub-carriers and $2M$ time slots as

$$x_i(n) = \sum_{k=0}^{K-1}\sum_{m=0}^{2M-1} j^{k+m} d_i^{k,m} g_m(n) e^{j\frac{2\pi kn}{K}},\ n = 0, 1, \cdots, N-1 \tag{1}$$

where $1 \le i \le N_b$ and N_b is the number of blocks per one transmission frame, $g_m(n) = g\left(n - m\frac{K}{2}\right)_N$ is a cyclic shift version of a length-N real-valued and symmetric prototype filter $g_0(n) = g(n)$. Equation (1) can be rewritten as

$$\mathbf{x}_i = \mathbf{A}\mathbf{d}_i \tag{2}$$

in which $\mathbf{d}_i = \left[\mathbf{d}_i^0\ \mathbf{d}_i^1\ \cdots\ \mathbf{d}_i^{K-1}\right]^T$ and $\mathbf{d}_i^q = \left[d_i^{q,0}\ d_i^{q,1}\ \cdots\ d_i^{q,2M-1}\right]^T$. To enable offset QAM, the real and imaginary parts of a complex QAM symbol in the ith transmission block, $s_i^{k,m} = s_{i,R}^{k,m} + j s_{i,I}^{k,m}$, are separated and arranged in (3), as shown at the bottom of this page. The real-valued vector $\mathbf{d}_i$ is then obtained by vectorizing the matrix on the left of (3). $\mathbf{A}$ is called the modulation matrix whose $(m + 2kM)$th column is $\mathbf{A}[:, m + 2kM] = j^{k+m}\left[g_m(0)e^{j\frac{2\pi k0}{K}}, \cdots, g_m(N-1)e^{j\frac{2\pi k(N-1)}{K}}\right]^T$ [15]. A CP of length L is added to $\mathbf{x}_i$ before transmitting.

The channel $\mathbf{h} = \left[h(0)\ h(1)\ \cdots\ h(\mu-1)\right]^T$ is characterized by μ taps, where $\mu \le (L+1)$. It is assumed that $\mathbf{h} \sim \mathcal{CN}(0, \mathbf{\Sigma})$ and $\mathbf{\Sigma} = \mathrm{E}\left[\mathbf{h}\mathbf{h}^H\right]$ is the $\mu \times \mu$ channel covariance matrix. When there is no PN, the received signal corresponding to the ith block after removing the CP is written as

$$\mathbf{y}_i = \mathbf{H}_{\text{circ}}\mathbf{x}_i + \mathbf{w}_i \tag{4}$$

where $\mathbf{w}_i \sim \mathcal{CN}\left(0, \sigma_{\mathbf{w}}^2\mathbf{I}\right)$ represents noise, and $\mathbf{H}_{\text{circ}}$ is a circulant matrix whose columns are circularly shifted from vector $\left[h(0)\ h(1)\ \cdots\ h(\mu-1)\ 0\ \cdots\ 0\right]_{N\times 1}^T$. It is noted that $\mathbf{H}_{\text{circ}}$ can be represented as $\mathbf{H}_{\text{circ}} = \mathbf{F}^H\mathbf{\Gamma}\mathbf{F}$ where $\mathbf{F}$ is the

$$\begin{bmatrix} d_i^{0,0} & d_i^{0,1} & \cdots & d_i^{0,2M-1} \\ d_i^{1,0} & d_i^{1,1} & \cdots & d_i^{1,2M-1} \\ \vdots & \vdots & \ddots & \vdots \\ d_i^{K-1,0} & d_i^{K-1,1} & \cdots & d_i^{K-1,2M-1} \end{bmatrix} = \begin{bmatrix} s_{i,R}^{0,0} & s_{i,I}^{0,0} & \cdots & s_{i,R}^{0,M-1} & s_{i,I}^{0,M-1} \\ s_{i,R}^{1,0} & s_{i,I}^{1,0} & \cdots & s_{i,R}^{1,M-1} & s_{i,I}^{1,M-1} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ s_{i,R}^{K-1,0} & s_{i,I}^{K-1,0} & \cdots & s_{i,R}^{K-1,M-1} & s_{i,I}^{K-1,M-1} \end{bmatrix}. \tag{3}$$

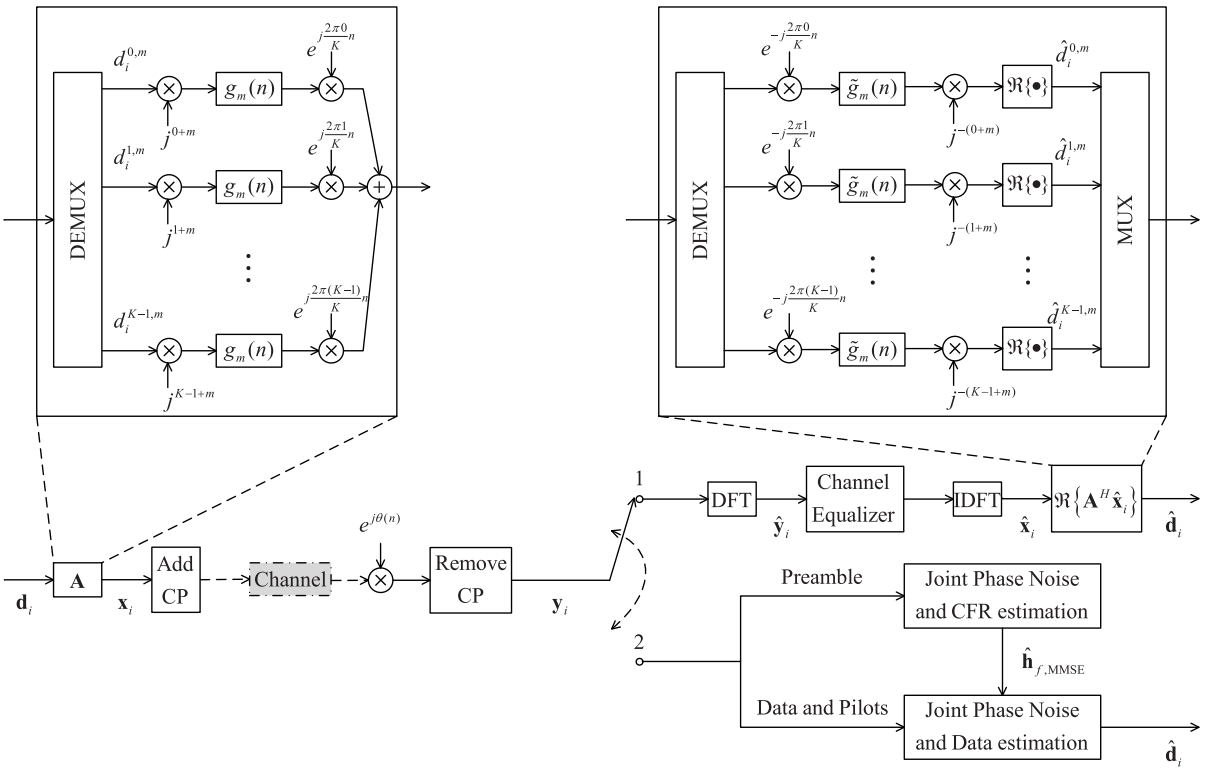


FIGURE 1. A discrete-time complex baseband equivalent CFBMC-OQAM transceiver.

normalized DFT matrix whose (k, m)th element is given as $[\mathbf{F}]_{k,m} = \frac{1}{\sqrt{N}} e^{-j\frac{2\pi mk}{N}}$ and $\mathbf{\Gamma}$ is a diagonal matrix whose diagonal elements are subcarrier channel gains in the frequency domain. To detect the desired signal, the received signal $\mathbf{y}_i$ is first converted to the frequency domain by a DFT transform, then channel equalized with matrix $\hat{\mathbf{\Gamma}}$ which is an estimate of $\mathbf{\Gamma}$ based on a known preamble, and further IDFT transformed to obtain $\hat{\mathbf{x}}_i$. Then, $\hat{\mathbf{x}}_i$ is passed through a bank of matched filters and the estimate of $\mathbf{d}_i$ is obtained by taking the real part of the filtered output as $\hat{\mathbf{d}}_i = \Re\left\{\mathbf{A}^H \hat{\mathbf{x}}_i\right\}$ [15].

In the presence of PN, the received signal for the ith block is

$$\mathbf{y}_i = \mathbf{P}_i \mathbf{H}_{\text{circ}} \mathbf{x}_i + \mathbf{w}_i \tag{5}$$

where $\mathbf{P}_i = \text{diag}\left[e^{j\theta(\epsilon)}\ e^{j\theta(\epsilon+1)} \cdots e^{j\theta(\epsilon+N-1)}\right]$ represents the effect of PN on the ith transmission block and $\epsilon = (i-1)N + iL$. It is pointed out that the form of the input/output relation as in (4) and (5) is common to many multicarrier systems, including OFDM, GFDM and CFBMC-OQAM, which use a CP to overcome the ISI effect of a frequency-selective fading channel. However, different signal processing operations result in either an orthogonal system (such as OFDM), or a non-orthogonal system (such as GFDM), or a real-domain orthogonal system (such as CFBMC-OQAM). Therefore, it is necessary to study different systems in detail when the impacts of PN and imperfect channel estimation are considered.

B. PHASE NOISE

This section briefly reviews two PN models: one is based on a free-running oscillator and the other is based on a PLL oscillator.

1) FR-PN MODEL

Let $\alpha_v(n)$ denote the discrete-time phase deviation from an FR oscillator at the nth sampling time. It can be modeled as $\alpha_v(n) = \sum_{i=0}^{n-1} \rho(i)$ where $\rho(i)$'s are independent and identically distributed (i.i.d) zero-mean Gaussian random variables with variance $\sigma_\rho^2 = c_v T_s$. Here, c_v is a constant describing the quality of an oscillator, and T_s is the sampling interval. The PN is $\theta(n) = 2\pi f_c \alpha_v(n) = \omega_c \alpha_v(n)$, where f_c is the carrier frequency. Thus, the discrete-time FR-PN at the nth sampling time $\theta(n)$ can be modeled as a zero-mean Gaussian distributed random variable with variance $n\omega_c^2 c_v T_s$, which grows linearly with the sample index n. Furthermore, the autocorrelation function of the phase deviation can be computed as $\mathrm{E}\left[\alpha_v(n_1)\alpha_v(n_2)\right] = c_v T_s \min(n_1, n_2)$ [44]. Define $\mathbf{R}$ as the PN covariance matrix, whose (n_1, n_2)th component is [25]:

$$\mathbf{R}(n_1, n_2) = \exp\left\{-\frac{1}{2}\left|n_1 - n_2\right| \omega_c^2 c_v T_s\right\}. \tag{6}$$

The above PN covariance matrix will be used later to compensate the impact of PN.

2) PLL-PN MODEL

A general PLL-based frequency synthesizer is described in [25]. Let $\alpha_r(n)$, $\alpha_v(n)$, and $\beta(n)$ denote the discrete-time phase deviations at the output of the reference oscillator, the frequency synthesizer, and the phase detector, respectively. Then $\alpha_v(n) = \beta(n) + \alpha_r(n)$ [45]. The reference oscillator is basically the same as a FR oscillator which is characterized by a quality factor c_r. The phase deviation output of the phase detector, i.e., $\beta(n)$, is modeled as a one-dimensional Ornstein-Uhlenbeck process [45]. Furthermore, the correlation properties between $\alpha_r(n)$ and $\beta(n)$ are: $\mathrm{E}\{\beta(n_1)\alpha_r(n_2)\} = \sum_{i=1}^{n_o} \mu_i \mathrm{e}^{\lambda_i T_s \min(0,n_2-n_1)}$ and $\mathrm{E}\{\beta(n_1)\beta(n_2)\} = \sum_{i=1}^{n_o} \nu_i \mathrm{e}^{-\lambda_i T_s |n_1-n_2|}$ where $n_o = 1 + o_{\mathrm{lpf}}$, and o_{lpf} represents the order of the loop filter [25], [45]. The values of λ_i, μ_i and ν_i are given, for instance in [45], for the PLL-based oscillator with a first-order loop filter. The (n_1, n_2)th component of the covariance matrix $\mathbf{R}$ is obtained as [25]

$$\mathbf{R}(n_1,n_2) = \exp\left\{ -\frac{1}{2}\omega_c^2 \left[|n_1-n_2| c_r T_s + 2\sum_{i=1}^{n_o} (\nu_i+\mu_i)\left(1-\mathrm{e}^{-\lambda_i T_s |n_1-n_2|}\right)\right]\right\}. \quad (7)$$

III. ESTIMATION OF PHASE NOISE AND CHANNEL FREQUENCY RESPONSE

It is clear that the estimated channel $\hat{\mathbf{\Gamma}}$ plays a critical role in detecting the transmitted signal. In practice, the channel state information can be obtained based on the transmission of preambles or pilot symbols. Specifically, preambles are training blocks which are transmitted at the beginning of each transmission frame, while pilots are known symbols which are embedded into transmission blocks based on a specific pattern and transmitted along with the data. Given known preambles or pilots, channel frequency response can be estimated by dividing the received signal by the corresponding preambles or pilots. Since a preamble occupies the entire transmission block, the channel estimation based on a preamble, if well designed, should be better than that obtained based on pilot symbols. Hence, a preamble is designed and used to estimate the channel in this paper, while pilot symbols are used to compensate for the PN.

Generally, channel estimation can be done in either time or frequency domain [4], [46], [47]. It appears that the frequency-domain approach is more common in practice for a multicarrier system since it can make use of the IFFT/FFT blocks already existing in the system and it also integrates well with the data detection process, which is conveniently performed in the frequency domain. As such, frequency-domain channel estimation is also adopted in this paper.

In this paper, it is assumed that the channel changes slowly during multiple transmission blocks, while PN changes fast and varies from one block to the other. Under this assumption the channel estimation based on the preamble, which is the first transmitted block in a transmission frame, can be used to compensate for the impact of PN and detect the data in the remaining transmission blocks. Thus, this paper proposes an algorithm which has two stages. Specifically, the channel frequency response and PN are estimated in the first stage based on the preamble. Given the estimated channel, data and PN can be estimated based on the pilot symbols in the second stage. To perform this two-stage PN compensation algorithm, the switch in Fig. 1 is toggled between the two positions.

A. CHANNEL ESTIMATION AND PREAMBLE DESIGN WITHOUT THE PRESENCE OF PN IN CFBMC-OQAM

First, the received signal corresponding to the first transmission block $\mathbf{y}_1$ is converted to the frequency domain by a DFT transform to as

$$\hat{\mathbf{y}}_1 = \mathbf{F}\mathbf{y}_1 = \mathbf{F}\mathbf{H}_{\mathrm{circ}}\mathbf{A}\mathbf{d}_1 + \mathbf{F}\mathbf{w}_1 = \mathbf{\Gamma}\mathbf{F}\mathbf{A}\mathbf{d}_1 + \hat{\mathbf{w}}_1. \quad (8)$$

Define $\mathbf{S}_1 = \mathrm{diag}[\mathbf{F}\mathbf{A}\mathbf{d}_1]$ and $\mathbf{h}_f = \mathrm{diag}[\mathbf{\Gamma}]$. Then (8) can be rewritten as

$$\hat{\mathbf{y}}_1 = \mathbf{S}_1\mathbf{h}_f + \hat{\mathbf{w}}_1. \quad (9)$$

For channel estimation, a known preamble is transmitted and hence $\mathbf{S}_1$ is known. When the LS estimation method is employed, the estimated channel frequency response $\hat{\mathbf{h}}_{f,\mathrm{LS}}$ is obtained by minimizing the following cost function:

$$C(\hat{\mathbf{h}}_{f,\mathrm{LS}}) = \left|\left|\hat{\mathbf{y}}_1 - \mathbf{S}_1\hat{\mathbf{h}}_{f,\mathrm{LS}}\right|\right|_2^2. \quad (10)$$

By setting the derivative of the cost function with respect to $\hat{\mathbf{h}}_{f,\mathrm{LS}}$ to zero, one obtains

$$\hat{\mathbf{h}}_{f,\mathrm{LS}} = \mathbf{S}_1^{-1}\hat{\mathbf{y}}_1. \quad (11)$$

It then follows that the MSE between $\hat{\mathbf{h}}_{f,\mathrm{LS}}$ and $\mathbf{h}_f$ is

$$\begin{aligned}
\mathrm{E}[||\hat{\mathbf{h}}_{f,\mathrm{LS}} - \mathbf{h}_f||_2^2] &= \mathrm{E}\left[\mathrm{Tr}\left((\hat{\mathbf{h}}_{f,\mathrm{LS}} - \mathbf{h}_f)(\hat{\mathbf{h}}_{f,\mathrm{LS}} - \mathbf{h}_f)^H\right)\right] \\
&= \mathrm{Tr}\left[\mathrm{E}\left(\mathbf{S}_1^{-1}\hat{\mathbf{y}}_1 - \mathbf{h}_f\right)\left(\mathbf{S}_1^{-1}\hat{\mathbf{y}}_1 - \mathbf{h}_f\right)^H\right] \\
&= \mathrm{Tr}\left[\mathrm{E}\left(\mathbf{S}_1^{-1}\left(\mathbf{S}_1\mathbf{h}_f + \hat{\mathbf{w}}_1\right) - \mathbf{h}_f\right) \times \right. \\
&\qquad \left.\left(\mathbf{S}_1^{-1}\left(\mathbf{S}_1\mathbf{h}_f + \hat{\mathbf{w}}_1\right) - \mathbf{h}_f\right)^H\right] \\
&= \mathrm{Tr}\left[\mathrm{E}\left(\mathbf{S}_1^{-1}\hat{\mathbf{w}}_1\hat{\mathbf{w}}_1^H\left(\mathbf{S}_1^{-1}\right)^H\right)\right] \\
&= \sigma_{\mathbf{w}}^2 \mathrm{Tr}\left\{\left(\mathbf{S}_1^H\mathbf{S}_1\right)^{-1}\right\}. \quad (12)
\end{aligned}$$

As expected the MSE depends on the preamble through the diagonal matrix $\mathbf{S}_1$. This also means that one can design the preamble to minimize the MSE.

Define $\mathbf{\Psi} = \mathbf{S}_1^H\mathbf{S}_1 = \mathrm{diag}\left[\psi_0\ \psi_1\ \cdots\ \psi_{N-1}\right]$. Then the preamble design problem is stated as

$$\begin{aligned}
&\underset{\psi_i}{\text{minimize}} \sum_{i=0}^{N-1} \frac{1}{\psi_i} \\
&\text{subject to } \psi_i \geq 0 \quad i = 0, 1, \cdots, N-1 \\
&\qquad \sum_{i=0}^{N-1} \psi_i \leq N. \quad (13)
\end{aligned}$$

In (13), the first constraint is to guarantee that $\mathbf{\Psi}$ is a positive semi-definite matrix, while the second constraint is to make sure that the total power of $\mathbf{S}_1$ is no more than N. The solution for the optimization problem in (13) can be obtained by applying the Karush-Kuhn-Tucker conditions [48] and the solution is $\mathbf{\Psi} = \mathbf{I}$. Thus, $\mathbf{S}_1$ is a diagonal matrix whose diagonal elements are in the forms of exponential functions, i.e.,

$$\mathbf{S}_1 = \text{diag}\left[e^{j\kappa_0}\ e^{j\kappa_1}\ \cdots\ e^{j\kappa_{N-1}}\right]^T. \tag{14}$$

Given $\mathbf{S}_1 = \text{diag}\left(\mathbf{F}\mathbf{A}\mathbf{d}_1\right)$ and the fact that $\Re\left\{\mathbf{A}^H\mathbf{A}\right\} = \mathbf{I}$ for a well-designed shaping filter [15], the preamble is finally obtained as

$$\mathbf{d}_1 = \Re\left\{\mathbf{A}^H\mathbf{F}^H\text{diag}\left[\mathbf{S}_1\right]\right\}. \tag{15}$$

The above preamble design can be straightforwardly applied to OFDM systems where the modulation matrix $\mathbf{A}$ is replaced by the inverse DFT matrix $\mathbf{F}^H$. This also means that to minimize the channel MSE, the preamble for OFDM under the LS channel estimation has to satisfy the unit-magnitude condition, namely $|\text{diag}[\mathbf{S}_1]| = \mathbf{1}_{N\times 1}$. This is consistent with previous studies for OFDM [16], [18], [49], [50].

It is pointed out that the preamble design in this paper does not take into account either the impact of PN or fading channel. This is a common practice (see e.g., [51], [52]) and is motivated by the fact that preambles are usually used for various purposes, such as establishing synchronization between the transmitter and receiver, estimating system parameters and channel statistics. To be applied ubiquitously, these preambles should be designed or selected without taking into account specific propagation channels and physical impairments. Given the designed preamble, the channel can be estimated based on either LS or MMSE method depending on whether statistical characteristics of channel and noise are available.

B. ESTIMATION OF PN AND CHANNEL FREQUENCY RESPONSE

In the presence of PN, this section is performed on the first transmission block, i.e., the designed preamble. Due to PN, the received signal (after removing CP) corresponding to the first transmission block becomes

$$\mathbf{y}_1 = \mathbf{P}_1\mathbf{H}_{\text{circ}}\mathbf{x}_1 + \mathbf{w}_1 = \mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f + \mathbf{w}_1 \tag{16}$$

where $\mathbf{P}_1 = \text{diag}\left[e^{j\theta(L)}\ e^{j\theta(L+1)}\ \cdots\ e^{j\theta(N+L-1)}\right]$ represents the effect of PN on the first transmission block.

It is seen from (16) that there exists an inter-dependency between PN and the channel, i.e., estimation of the channel depends on PN and vice versa. To estimate PN and the channel, the approach in [31], [32], [34] is adopted. In such an approach, the inter-dependency is decoupled by first obtaining the LS-based channel estimate which is a function of the unknown PN matrix. Then, the estimated channel is substituted into the optimization problem which minimizes the MSE for PN estimation. However, different from the previous studies, the MMSE method is used to estimate the channel in this paper. Compared to LS-based channel estimation, the MMSE method takes into account statistical characteristics of both channel and noise to improve the channel estimation quality. To focus on the PN compensation algorithm, it is assumed that the statistical characteristics of channel and noise are known at the receiver.[2]

The MMSE channel estimate $\hat{\mathbf{h}}_{f,\text{MMSE}}$ is found by minimizing $\text{E}\left[||\mathbf{h}_f - \hat{\mathbf{h}}_{f,\text{MMSE}}||_2^2\right]$ [53]. The orthogonality principle states that

$$\text{E}\left[\left(\mathbf{h}_f - \hat{\mathbf{h}}_{f,\text{MMSE}}\right)\hat{\mathbf{h}}_{f,\text{MMSE}}^H\right] = \mathbf{0}, \tag{17}$$

which results in

$$\hat{\mathbf{h}}_{f,\text{MMSE}} = N\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H \times\left(N\mathbf{S}_1\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F}\mathbf{P}_1^H\mathbf{y}_1 \tag{18}$$

where $\breve{\mathbf{F}}$ contains the first μ columns of $\mathbf{F}$. A proof for (18) is given in Appendix A. Recall that $\hat{\mathbf{h}}_{f,\text{LS}} = \left(\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\right)^\dagger\mathbf{y}_1 = \mathbf{S}_1^{-1}\mathbf{F}\mathbf{P}_1^H\mathbf{y}_1$. Thus, it can be seen from (18) that the MMSE channel estimate can be expressed as a linearly-weighted version of the LS channel estimate, namely $\hat{\mathbf{h}}_{f,\text{MMSE}} = \mathbf{W}\hat{\mathbf{h}}_{f,\text{LS}}$, where

$$\mathbf{W} = N\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H\left(N\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H + \sigma_{\mathbf{w}}^2\left(\mathbf{S}_1^H\mathbf{S}_1\right)^{-1}\right)^{-1}. \tag{19}$$

With the expression of the MMSE channel estimate $\hat{\mathbf{h}}_{f,\text{MMSE}}$, PN is then estimated to minimize the following MSE:

$$\mathbf{P}_1 = \arg\min_{\mathbf{P}_1}\left|\left|\left(\mathbf{y}_1 - \mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f\right)\Big|_{\mathbf{h}_f=\hat{\mathbf{h}}_{f,\text{MMSE}}}\right|\right|_2^2. \tag{20}$$

Let $\mathbf{B} = N\mathbf{F}^H\mathbf{S}_1\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H\left(N\mathbf{S}_1\breve{\mathbf{F}}\mathbf{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F}$. Then the cost function in (20) can be rewritten as

$$\begin{aligned} C(\mathbf{P}_1) &= \left|\left|\left(\mathbf{y}_1 - \mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f\right)\Big|_{\mathbf{h}_f=\hat{\mathbf{h}}_{f,\text{MMSE}}}\right|\right|_2^2 \\ &= \left|\left|\mathbf{y}_1 - \mathbf{P}_1\mathbf{B}\mathbf{P}_1^H\mathbf{y}_1\right|\right|_2^2 \\ &= \mathbf{y}_1^H\mathbf{y}_1 - \mathbf{y}_1^H\mathbf{P}_1\left(\mathbf{B} + \mathbf{B}^H\right)\mathbf{P}_1^H\mathbf{y}_1 \\ &\quad + \mathbf{y}_1^H\mathbf{P}_1\mathbf{B}^H\mathbf{B}\mathbf{P}_1^H\mathbf{y}_1 \\ &= \mathbf{y}_1^H\mathbf{P}_1\left(\mathbf{I} - 2\mathbf{B} + \mathbf{B}^H\mathbf{B}\right)\mathbf{P}_1^H\mathbf{y}_1. \end{aligned} \tag{21}$$

The last equation in (21) is derived based on the fact that $\mathbf{P}_1\mathbf{P}_1^H = \mathbf{I}$ and $\mathbf{B}$ is a Hermitian matrix, i.e., $\mathbf{B} = \mathbf{B}^H$. Thus, the optimization problem (20) becomes

$$\mathbf{P}_1 = \arg\min_{\mathbf{P}_1}\mathbf{y}_1^H\mathbf{P}_1\left(\mathbf{I} - 2\mathbf{B} + \mathbf{B}^H\mathbf{B}\right)\mathbf{P}_1^H\mathbf{y}_1. \tag{22}$$

[2]In practice, statistical information of channel and noise can be obtained using either data-aided techniques (which require and make use of pilot symbols) or non-data aided techniques (which are based directly on the received signals).

Define $\mathbf{U} = (\mathbf{I} - 2\mathbf{B} + \mathbf{B}^H\mathbf{B})$, which is a Hermitian matrix, $\mathbf{Y}_1 = \text{diag}\,[\mathbf{y}_1]$, and $\boldsymbol{\rho} = \begin{bmatrix}\rho_0 & \rho_1 & \cdots & \rho_{N-1}\end{bmatrix}^T = \text{diag}\,[\mathbf{P}_1]^H$. Then the optimization problem (22) can be rewritten as

$$\begin{aligned} &\underset{\boldsymbol{\rho}}{\text{minimize}}\ \boldsymbol{\rho}^H\mathbf{Y}_1^H\mathbf{U}\mathbf{Y}_1\boldsymbol{\rho} \\ &\text{subject to } |\rho_i| = 1; \quad i = 0, 1, \cdots, N-1. \end{aligned} \tag{23}$$

A similar form of the optimization problem (23) can be found in [31], [32], [34]. Similar to [34], the majorization-minimization algorithm is applied to solve (23). The principle behind the majorization-minimization algorithm is to transform a difficult problem into a series of simpler problems. This algorithm consists of two steps. In the first step, i.e., majorization, a surrogate function needs to be found which locally approximates the objective function at the current point. In the second step, i.e., minimization, the surrogate function is minimized.

From (23), define $\mathbf{V} = \mu_{\max}\mathbf{I}_{N\times N}$ where $\mu_{\max}$ is the maximum eigenvalue of $\mathbf{U}$. The surrogate function corresponding to the objective function in (23) is [38]

$$\begin{aligned} g(\boldsymbol{\rho}) = \boldsymbol{\rho}^H\mathbf{Y}_1^H\mathbf{V}\mathbf{Y}_1\boldsymbol{\rho} + 2\Re(\boldsymbol{\rho}^H\mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t) \\ +\boldsymbol{\rho}_t^H\mathbf{Y}_1^H(\mathbf{V}-\mathbf{U})\mathbf{Y}_1\boldsymbol{\rho}_t \end{aligned} \tag{24}$$

in which $\boldsymbol{\rho}_t$ is an arbitrary vector. The first and third terms in (24) are constants and independent of $\boldsymbol{\rho}$. Thus, minimizing $g(\boldsymbol{\rho})$ in the second step is equivalent to minimizing the second term in (24). Hence the optimization problem (23) becomes

$$\begin{aligned} &\underset{\boldsymbol{\rho}}{\text{minimize}}\ \Re(\boldsymbol{\rho}^H\mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t) \\ &\text{subject to } |\rho_i| = 1; \quad i = 0, 1, \cdots, N-1. \end{aligned} \tag{25}$$

A closed-form solution for (25) is obtained as:

$$\boldsymbol{\rho}_{t+1} = \mathrm{e}^{j\arg(\mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t)}. \tag{26}$$

A proof for (26) is given in Appendix B.

In summary, to find a solution for the optimization problem (23), $\boldsymbol{\rho}_t$ is first initialized at $t = 0$, i.e., $\boldsymbol{\rho}_0$, and then substituted into (26) to obtain $\boldsymbol{\rho}_1$. Next, $\boldsymbol{\rho}_1$ is substituted into (26) to obtain $\boldsymbol{\rho}_2$ and so on. The algorithm terminates after a predefined number of iterations N_i or when the squared ℓ_2-norm of the difference over two consecutive iterations is smaller than a predefined threshold. The estimated PN is then substituted into (18) to finally obtain the estimated channel frequency response.

The computational complexity of the first stage is proportional to the number of iterations performed on (26). It is pointed out that $\mathbf{G} = \mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1$ in (26) is a constant and is computed once before any iterations. Thus, for each iteration, the computational complexity is determined by the multiplication between $\mathbf{G}$ and $\boldsymbol{\rho}_t$, which requires N^2 complex multiplications and $N(N-1)$ complex additions. For other parameters, they are also computed once before any iterations. Furthermore, their computational complexities can be reduced by utilizing the facts that: (i) the first stage is performed based on the transmission of the designed preamble which satisfies the condition that $\mathbf{S}_1$ is a diagonal matrix and $\mathbf{S}_1\mathbf{S}_1^H = \mathbf{S}_1^H\mathbf{S}_1 = \mathbf{I}$, and (ii) the statistical characteristics of the channel and noise do not change much from one frame, hence $\boldsymbol{\Sigma}$ and $\sigma_{\mathbf{w}}$ can be considered as constants.

For example, matrix $\mathbf{B}$ can be rewritten as

$$\begin{aligned} \mathbf{B} &= N\mathbf{F}^H\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\left(N\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F} \\ &\overset{(40)}{=} N\mathbf{F}^H\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\left(N\check{\mathbf{F}}^H\check{\mathbf{F}}\boldsymbol{\Sigma} + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\check{\mathbf{F}}^H\mathbf{S}_1^H\mathbf{F} \\ &\overset{(40)}{=} \mathbf{F}^H\mathbf{S}_1\mathbf{T}\mathbf{S}_1^H\mathbf{F} \end{aligned} \tag{27}$$

where $\mathbf{T} = N\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\left(N\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}$. In the above equation, the computational complexity due to matrix inversion is significantly reduced from inverting an $N \times N$ matrix to inverting a $\mu \times \mu$ matrix, where $\mu \ll N$. The matrix $\mathbf{T}$ is constant and does not depend on the transmitted preamble. Only the DFT of the preamble, i.e., $\mathbf{F}^H\mathbf{S}_1$, needs to be computed for every frame to update $\mathbf{B}$. Similarly, the computation of $\mathbf{B}^H\mathbf{B}$ is

$$\mathbf{B}^H\mathbf{B} = \mathbf{F}^H\mathbf{S}_1\mathbf{T}^2\mathbf{S}_1^H\mathbf{F} \tag{28}$$

which requires the update of $\mathbf{F}^H\mathbf{S}_1$ for every frame only.

IV. PHASE NOISE COMPENSATION AND DATA DETECTION

In this part, the MMSE-based estimated channel obtained from the previous section shall be used together with pilot symbols to compensate for PN and detect the transmitted signal in the data transmission phase. To estimate and compensate for PN, the PN vector shall be represented using known basis vectors. To this end, define $\boldsymbol{\Xi}_i = \begin{bmatrix}\boldsymbol{\xi}_i^0 & \boldsymbol{\xi}_i^1 & \cdots & \boldsymbol{\xi}_i^{F-1}\end{bmatrix}$, which contains F basis vectors, in which $\boldsymbol{\xi}_i^f = \begin{bmatrix}\xi_i^{f,0} & \xi_i^{f,1} & \cdots & \xi_i^{f,N-1}\end{bmatrix}^T$. The F basis vectors are chosen either from N columns of the DFT matrix or from N eigenvectors corresponding to the F largest eigenvalues of the PN covariance matrices as given in (6) and (7) for FR PN and PLL PN, respectively. Also define $\boldsymbol{\gamma}_i = \begin{bmatrix}\gamma_{i,0} & \gamma_{i,1} & \cdots & \gamma_{i,F-1}\end{bmatrix}^T$. In order to compensate for the PN, the goal is to estimate $\boldsymbol{\gamma}_i$ such that $\text{diag}\left[\boldsymbol{\Xi}_i\boldsymbol{\gamma}_i\right] \simeq \mathbf{P}_i^H$. Then, $\boldsymbol{\Xi}_i\boldsymbol{\gamma}_i$ can be used to de-rotate the received signal in the time domain prior to the DFT transform.

Specifically, given the estimated channel frequency response, i.e., $\hat{\boldsymbol{\Gamma}} = \text{diag}[\hat{\mathbf{h}}_{f,\text{MMSE}}]$, the demodulated signal for the ith transmission block after PN compensation can be written as

$$\begin{aligned} \hat{\mathbf{d}}_i &= \Re\left\{\mathbf{A}^H\mathbf{F}^H\hat{\boldsymbol{\Gamma}}^{-1}\mathbf{F}\text{diag}\left[\mathbf{P}_i\mathbf{H}_{\text{circ}}\mathbf{A}\mathbf{d}_i + \mathbf{w}_i\right]\boldsymbol{\Xi}_i\boldsymbol{\gamma}_i\right\} \\ &= \Re\left\{\boldsymbol{\Delta}_i\boldsymbol{\gamma}_i\right\}. \end{aligned} \tag{29}$$

Denote $\mathbf{p} = \left\{p_0, p_1, \cdots, p_{Q-1}\right\}$ as a set of pilot indexes in each transmission block, and the transmitted signal corresponding to the pilot index set as $\mathbf{d}_i^{[\mathbf{p}]} = \begin{bmatrix}d_i^{p_0} & d_i^{p_1} & \cdots & d_i^{p_{Q-1}}\end{bmatrix}^T$. Then, $\boldsymbol{\gamma}_i$ can be estimated to minimize the MSE between the originally transmitted pilots and the

estimated pilots as

$$\boldsymbol{\gamma}_i = \arg\min_{\boldsymbol{\gamma}_i} \left|\left|\mathbf{d}_i^{[\mathbf{p}]} - \Re\left\{\boldsymbol{\Delta}_i^{[\mathbf{p}]}\boldsymbol{\gamma}_i\right\}\right|\right|_2^2 \tag{30}$$

where $\boldsymbol{\Delta}_i^{[\mathbf{p}]}$ is a $Q \times F$ matrix obtained by keeping Q rows of $\boldsymbol{\Delta}_i$ with respect to the pilot index set $\mathbf{p}$. The matrix $\Re\left\{\boldsymbol{\Delta}_i^{[\mathbf{p}]}\boldsymbol{\gamma}_i\right\}$ is expanded in (31), as shown at the bottom of this page, where $\delta_i^{k,j}$ denotes the (k,j)th element of matrix $\boldsymbol{\Delta}_i^{[\mathbf{p}]}$. Furthermore, vector $\boldsymbol{\gamma}_i$ is related to $\tilde{\boldsymbol{\gamma}}_i$ as

$$\boldsymbol{\gamma}_i = \begin{bmatrix} 1 & -j & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -j & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -j \end{bmatrix} \tilde{\boldsymbol{\gamma}}_i$$

where $\tilde{\boldsymbol{\gamma}}_i$ can be found by solving

$$\tilde{\boldsymbol{\gamma}}_i = \arg\min_{\tilde{\boldsymbol{\gamma}}_i} ||\mathbf{d}_i^{[\mathbf{p}]} - \hat{\boldsymbol{\Delta}}_i\tilde{\boldsymbol{\gamma}}_i||_2^2 \tag{32}$$

which yields $\tilde{\boldsymbol{\gamma}}_i = \hat{\boldsymbol{\Delta}}_i^{\dagger}\mathbf{d}_i^{[\mathbf{p}]}$. To avoid the singularity when performing the matrix inverse in $\hat{\boldsymbol{\Delta}}_i^{\dagger}$, the number of pilots should be greater than or equal to the number of bases.

Compared to the iterative PN compensation algorithm proposed in [35], where the PN estimation is based on a one-tap LMS algorithm, the second stage proposed in this paper has higher computational complexity due to the matrix multiplication and inversion in (32). It is noted, however, that the complexity in (32) only depends on the numbers of bases and pilots which are much smaller than N. As will be shown later, good PN compensation performance can be achieved even when the numbers of bases and pilots are not too large. Furthermore, compared to the iterative PN compensation algorithm in [35], the second stage in our proposed algorithm does not require any iterations. Thus, the proposed algorithm is expected to have lower latency than the one in [35] when a long transmission block is considered.

The proposed PN compensation algorithm is summarized as in Algorithm 1. It contains two stages. The estimation in the first stage is initialized with $\boldsymbol{\rho}_0 = \mathbf{1}$ in order to obtain the PN estimate in (26) after N_i iterations. Given the estimated PN, i.e., $\mathbf{P}_1 = \text{diag}[\boldsymbol{\rho}_{N_i}]^H$, the channel then can be estimated based on (18). In the second stage, given the estimated channel obtained from the first stage, i.e., $\hat{\boldsymbol{\Gamma}} = \text{diag}[\hat{\mathbf{h}}_{f,\text{MMSE}}]$, as well as the transmitted pilot symbols, which are determined by the pilot index set $\mathbf{p}$, the PN corresponding to the ith data transmission block, which is represented using a known basis matrix $\boldsymbol{\Xi}_i$ and an unknown coefficient vector $\boldsymbol{\gamma}_i$, can be estimated by obtaining an estimate of $\boldsymbol{\gamma}_i$ from (32). Given the estimated PN, the impact of PN can be compensated accordingly.

Algorithm 1 Two-Stage Phase Noise Compensation

// Stage 1:
$\mathbf{T} = N\mathbf{S}_1\breve{\mathbf{F}}\boldsymbol{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H$;
$\mathbf{B} = \mathbf{F}^H\mathbf{T}\left(\mathbf{T} + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F}$;
$\mathbf{U} = \mathbf{I} - 2\mathbf{B} + \mathbf{B}^H\mathbf{B}$;
$\mathbf{V} = \mu_{\max}\mathbf{I}$, $\mu_{\max}$ is the maximum eigenvalue of $\mathbf{U}$;
Initialize: $t \leftarrow 0$; $\boldsymbol{\rho}_t \leftarrow \mathbf{1}$;
while $(t < N_i)$ **do**
  $\boldsymbol{\rho}_{t+1} = \mathrm{e}^{j\arg(\mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t)}$;
  $t = t + 1$;
end while
return $\boldsymbol{\rho}_{N_i}$;
$\mathbf{P}_1 = \text{diag}[\boldsymbol{\rho}_{N_i}]^H$;
$\hat{\mathbf{h}}_{f,\text{MMSE}} = N\breve{\mathbf{F}}\boldsymbol{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H(\mathbf{T} + \sigma_{\mathbf{w}}^2\mathbf{I})^{-1}\mathbf{F}\mathbf{P}_1^H\mathbf{y}_1$;
// Stage 2:
$\hat{\boldsymbol{\Gamma}} = \text{diag}[\hat{\mathbf{h}}_{f,\text{MMSE}}]$;
for $i = 2$ *to* N_b **do**
  $\mathbf{p} \Rightarrow \mathbf{d}_i^{[\mathbf{p}]}$;
  $\boldsymbol{\Delta}_i = \mathbf{A}^H\mathbf{F}^H\hat{\boldsymbol{\Gamma}}^{-1}\mathbf{F}\text{diag}\left[\mathbf{y}_i\right]\boldsymbol{\Xi}_i$; $\boldsymbol{\Delta}_i \Rightarrow \hat{\boldsymbol{\Delta}}_i$;
  $\tilde{\boldsymbol{\gamma}}_i = \hat{\boldsymbol{\Delta}}_i^{\dagger}\mathbf{d}_i^{[\mathbf{p}]}$; $\tilde{\boldsymbol{\gamma}}_i \Rightarrow \boldsymbol{\gamma}_i$;
  $\mathbf{P}_i^H = \text{diag}\left[\boldsymbol{\Xi}_i\boldsymbol{\gamma}_i\right]$;
  $\hat{\mathbf{y}}_i = \mathbf{P}_i^H\mathbf{y}_i$;
  Decoding $\hat{\mathbf{y}}_i$;
end for

V. SIMULATION RESULTS

In this section, performance of the proposed two-stage PN compensation algorithm is evaluated in different scenarios and with parameters listed in Table 1. Among many shaping filters [8], the linear-phase square-root raised cosine (SRRC) and Martin [9], which are widely adopted in single-carrier data transmissions, can also be used as prototype filters in FBMC-based systems as long as the underlying channel is slowly varying. As such, the SRRC and Martin filters shall be considered in this paper.[3] Because of the difference in the signal types being transmitted in the preamble and data blocks, performance in this section is evaluated with respect to the

[3]It might be interesting to examine other prototype filters such as the extended Gaussian filter (EGF) [55], [56], which has better time-frequency localization property.

$$\Re\left\{\boldsymbol{\Delta}_i^{[\mathbf{p}]}\boldsymbol{\gamma}_i\right\} = \begin{bmatrix} \Re\left\{\delta_i^{0,0}\right\} & \Im\left\{\delta_i^{0,0}\right\} & \cdots & \Re\left\{\delta_i^{0,F-1}\right\} & \Im\left\{\delta_i^{0,F-1}\right\} \\ \Re\left\{\delta_i^{1,0}\right\} & \Im\left\{\delta_i^{1,0}\right\} & \cdots & \Re\left\{\delta_i^{1,F-1}\right\} & \Im\left\{\delta_i^{1,F-1}\right\} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \Re\left\{\delta_i^{Q-1,0}\right\} & \Im\left\{\delta_i^{Q-1,0}\right\} & \cdots & \Re\left\{\delta_i^{Q-1,F-1}\right\} & \Im\left\{\delta_i^{Q-1,F-1}\right\} \end{bmatrix} \begin{bmatrix} \Re\left\{\gamma_{i,0}\right\} \\ -\Im\left\{\gamma_{i,0}\right\} \\ \cdots \\ \Re\left\{\gamma_{i,F-1}\right\} \\ -\Im\left\{\gamma_{i,F-1}\right\} \end{bmatrix} = \hat{\boldsymbol{\Delta}}_i\tilde{\boldsymbol{\gamma}}_i \tag{31}$$

TABLE 1. Simulation parameters.

Channel coding	Convolutional code (Rate: 1/3; Constraint length: 7; Generator polynomials: $[133\ \ 171\ \ 165]_8$);
Channel decoding	Hard-decision Viterbi decoding;
Modulation	16-QAM;
Block structure	$K = 128$; $M = 6$;
Prototype filter	Martin; SRRC (Roll-off factor: 0.5);
Number of blocks per frame	$N_b = 10$;
Carrier frequency	$f_c = 6$ GHz; [39]
Bandwidth	$B = 100$ MHz;
Sampling interval	$T_s = 1/B$;
Channel [54]	EPA ($L = 41$); EVA ($L = 251$);
Phase noise [45]	$c_v = 5 \times 10^{-18}$; $c_r = 10^{-23}$

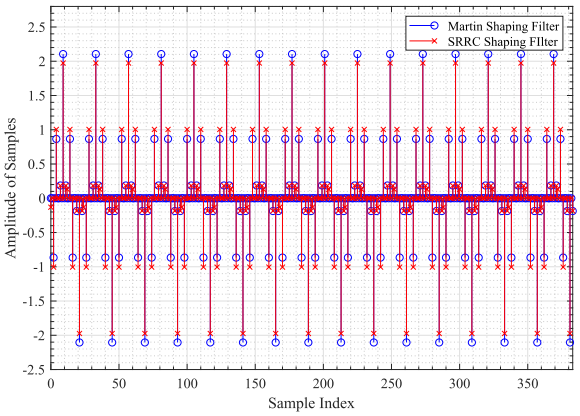


FIGURE 2. The designed preamble $\mathbf{d}_1$ when $K = 64$, $M = 3$, $\mathbf{S}_1 = \mathbf{I}_{N\times N}$, and Martin and SRRC shaping filters are deployed.

received signal-to-noise ratio (SNR). Two highly frequency-selective channel models, namely the extended pedestrian A (EPA) and extended vehicular A (EVA), are considered in this paper. In the transmitter, a CP length of $L = 41$ or $L = 251$ samples is added into the signal when EPA or EVA is considered, respectively.

A. PREAMBLE DESIGN FOR CHANNEL ESTIMATION

First, the preambles designed to mitigate the performance degradation in CFBMC-OQAM compared to OFDM due to imperfect channel estimation corresponding to SRRC and Martin filters are illustrated in Fig. 2. To intuitively show the difference between them, a short transmission block is applied where the numbers of sub-carriers and time slots are $K = 64$ and $2M = 6$, respectively. In this figure, $\mathbf{S}_1$ is simply chosen as an identity matrix to satisfy the condition (14). As such, the proposed preamble is obtained as

$$\mathbf{d}_1 = \Re\left\{\mathbf{A}^H\mathbf{F}^H\mathbf{1}_{N\times 1}\right\}. \tag{33}$$

In Fig. 2, it is observed that applying different prototype shaping filters results in different preamble patterns.

Performance comparisons of different preamble designs in terms of MSE and BER over EPA channel are plotted in Fig. 3 when there is no PN (i.e., $\mathbf{P}_1$ is replaced by an identity matrix in (18)). Performance curves of OFDM are included as benchmarks. In simulation, "Random Preamble" refers to the case that each symbol in the OFDM preamble is randomly generated from 16-QAM modulation, while each symbol in the CFBMC-OQAM preamble is either a real or an imaginary part of a randomly-generated 16-QAM symbol. On the other hand, "Proposed Preamble" means that the OFDM preamble contains randomly-generated BPSK symbols, while in CFBMC-OQAM, each element in $\mathbf{S}_1$ is a randomly-generated BPSK symbol to satisfy the unit-magnitude condition in (14). It is noted that OFDM preamble is a $N \times 1$ vector while it is a $2N \times 1$ vector for CFBMC-OQAM preamble. Also for simplicity, in presenting the following results, only the Martin filter is considered.

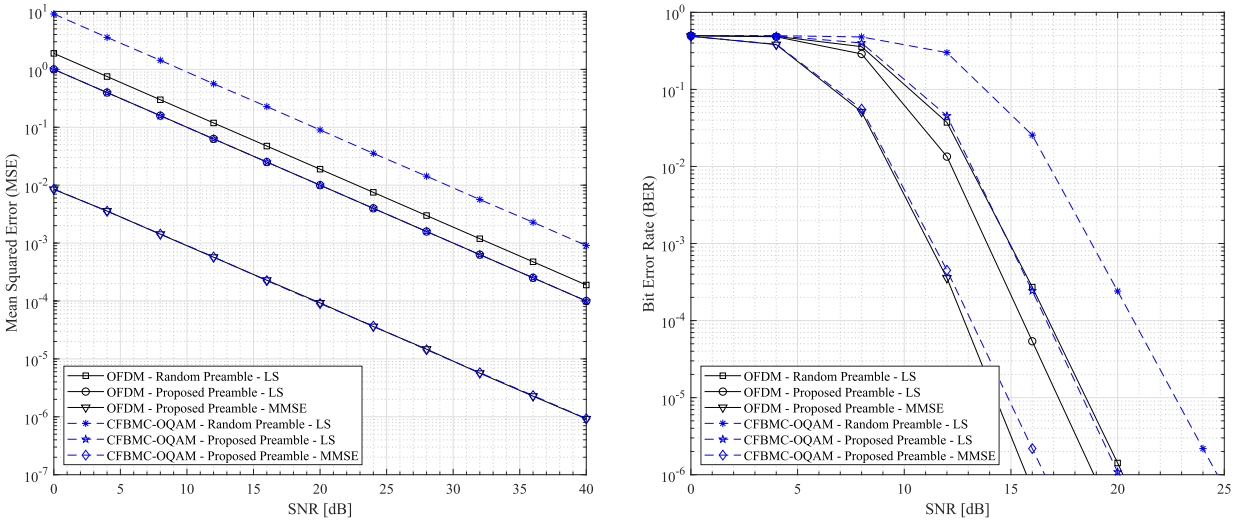
It is first observed from Fig. 3 that when a "Random Preamble" and LS estimation are deployed, there is a large performance degradation of about 6.8 dB at MSE $= 10^{-3}$ in CFBMC-OQAM compared to that in OFDM. While in terms of BER, the performance loss of CFBMC-OQAM is approximately 4.1 dB at BER $= 10^{-5}$ compared to that in OFDM under the same configuration. However, when the proposed preambles are applied in both OFDM and CFBMC-OQAM, the performance of CFBMC-OQAM in terms of MSE is almost the same as that in OFDM for both LS and MMSE channel estimation methods. For BER performance, there are only about 1.1 dB and 0.6 dB losses at BER $= 10^{-5}$ for CBMC-OQAM compared to OFDM when LS and MMSE methods are deployed, respectively.

For CFBMC-OQAM, the proposed preamble helps to gain approximately 9.6 dB and 30.0 dB at MSE $= 10^{-3}$ when LS and MMSE are used, respectively, compared to when a random preamble is deployed. While in terms of BER performance, these gains are around 4.2 dB and 7.9 dB at BER $= 10^{-5}$, respectively.

It is pointed out that at high SNR, performance of the MMSE channel estimator approaches that of the LS channel estimator, provided that these estimators operate in the time domain and that the channel taps are statistically independent (which is a common assumption in the literature). The behavior is, however, different for the MMSE and LS channel estimators operating in the frequency domain, as in this paper. This is because, as long as the number of subchannels in the frequency domain is larger than the number of channel taps in the time domain (which is typically the case in practice), the subchannels in the frequency domain are correlated. Since the MMSE estimator takes into account the correlation of the subchannels as well as the noise statistics, it consistently outperforms the LS estimator (which does not make use of the subchannels correlation or noise statistics). In fact, a similar performance gap (even at high SNR) between the frequency-domain MMSE and LS channel estimators is also observed for the OFDM systems in [46].

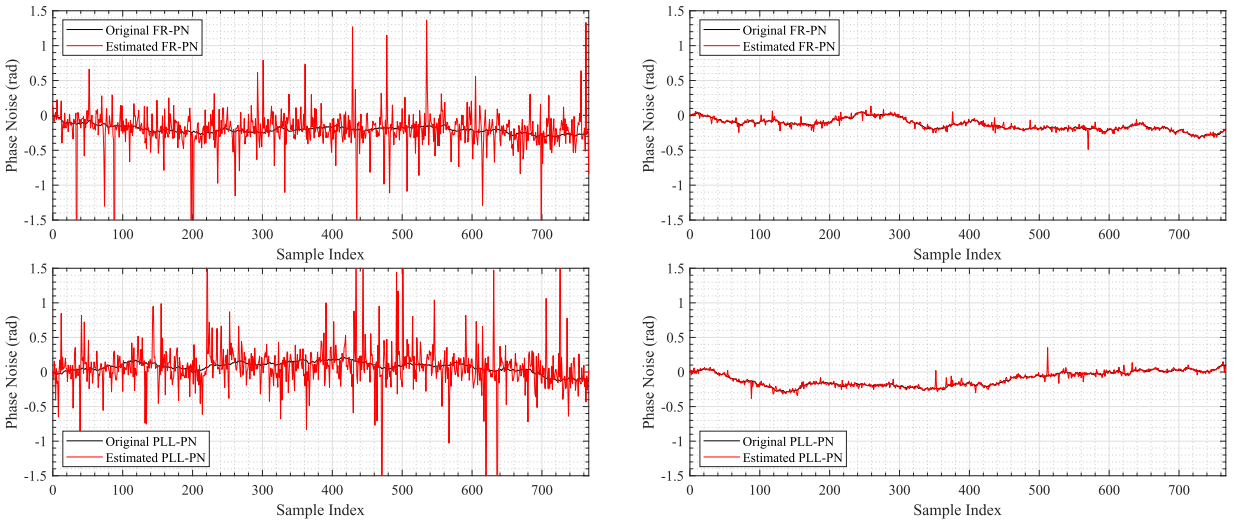
B. STAGE 1: ESTIMATION OF PN AND CHANNEL FREQUENCY RESPONSE

From this part, impact of PN is included in simulation results. The PN estimation algorithm in Section III-B is ini-



(a) MSE (b) BER

FIGURE 3. Performance comparisons in terms of (a) MSE and (b) BER for different preamble designs over EPA channel when $K = 128$, $M = 6$ and Martin shaping filter is deployed.



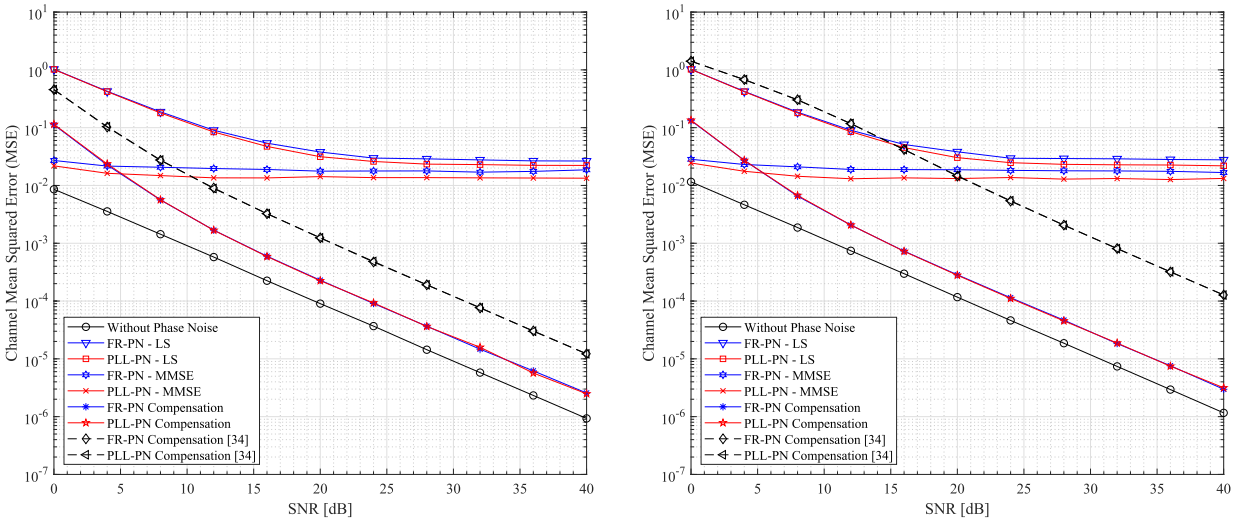
(a) SNR = 15 dB (b) SNR = 35 dB

FIGURE 4. Estimated PN in the first stage over EPA channel when $K = 128$, $M = 6$ and Martin shaping filter is deployed.

tialized with $\boldsymbol{\rho}_0 = \mathbf{1}_{N\times 1}$. The number of iterations is set to $N_i = 150$. First, the estimated PN from the first stage is illustrated in Fig. 4 for SNR = 15 dB and SNR = 35 dB for both types of PN. In this figure, EPA channel is considered. It can be seen that although SNR = 15 dB is quite low, the algorithm can generally track the PN, but there are large fluctuations in the estimated PN. However, when increasing SNR to 35 dB, the PN is estimated almost perfectly.

After obtaining the PN estimation in (26), the estimated PN is substituted into (18) to estimate the channel frequency response. An evaluation in terms of MSE for the channel estimation is demonstrated in Fig. 5 which contains two sub-figures 5a and 5b corresponding to EPA and EVA channels, respectively. In Fig. 5, the performance for the case when there is no PN and the MMSE-based channel estimate is employed (labeled as ''Without Phase Noise''), serves as the lower bound. The performance curves under the presence of two PN types without PN compensation deploying LS and MMSE channel estimations are also included. Performance of the proposed algorithm in [34] is also added for two PN types. It is recalled that different from [34] in which the channel is estimated based on LS method, the channel is estimated based on MMSE method in this paper.

It can be seen from Fig. 5 that without PN compensation the system basically fails in estimating the channel as the channel MSE is always above 10^{-2} for the whole range of SNR in



(a) EPA channel (b) EVA channel

FIGURE 5. Channel MSE from the first stage under two PN models when $K = 128$, $M = 6$ and Martin shaping filter is deployed.

both sub-figures. It is also observed that the channel MSE performance when PLL-PN is present and no compensation is applied is lower than that when FR-PN is present for both LS and MMSE channel estimation methods. This is expected because the impact of FR-PN is more significant than that of PLL-PN.

By applying the channel and PN estimation in the first stage of the two-stage algorithm, the channel MSE is significantly improved. Specifically, for EPA channel, there is only approximately 4.4 dB performance loss compared to the ideal case ("Without Phase Noise") at MSE $= 10^{-3}$. For EVA channel, this performance gap is around 4.1 dB at MSE $= 10^{-3}$. It can be observed that performances of the proposed channel estimation in the first stage are almost the same for two types of PN in both Figs. 5a and 5b.

The results in Fig. 5 show that, compared to the algorithm in [34], the proposed algorithm in this paper yields significant gains. Specifically, under the case of EPA channel, approximately 6.9 dB performance gain is achieved at MSE $= 10^{-3}$. This improvement is further increased to approximately 16.3 dB when the EVA channel is considered. It will be seen later that the performance gain brought by channel estimation in this stage significantly improves the PN compensation in the second stage of the proposed algorithm.

On the other hand, the degraded channel MSE in the low SNR observed from Fig. 5 is similar to the observation in [34]. This is caused by the susceptibility of the estimation in the first stage to the under-determination problem in the low SNR range, where the number of unknown variables is larger than the number of equations provided by the received signal.

It is interesting to see from Fig. 4a that, although there is a high fluctuation in the estimated PN compared to the real PN, the first stage still results in a better channel MSE performance compared to the case when no PN compensation is applied, i.e., by ignoring PN, as shown in Fig. 5. To explain

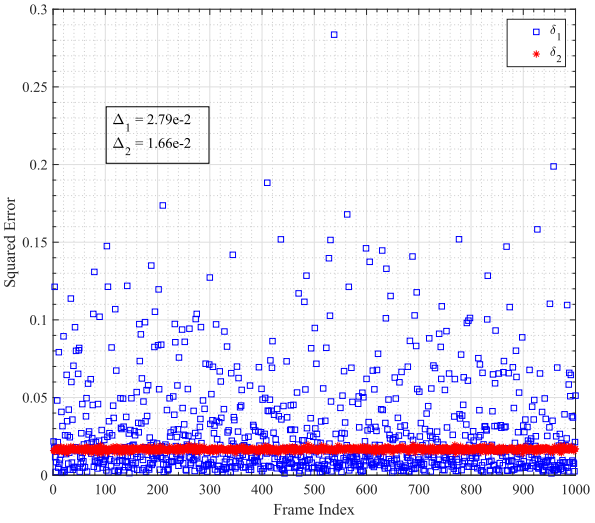
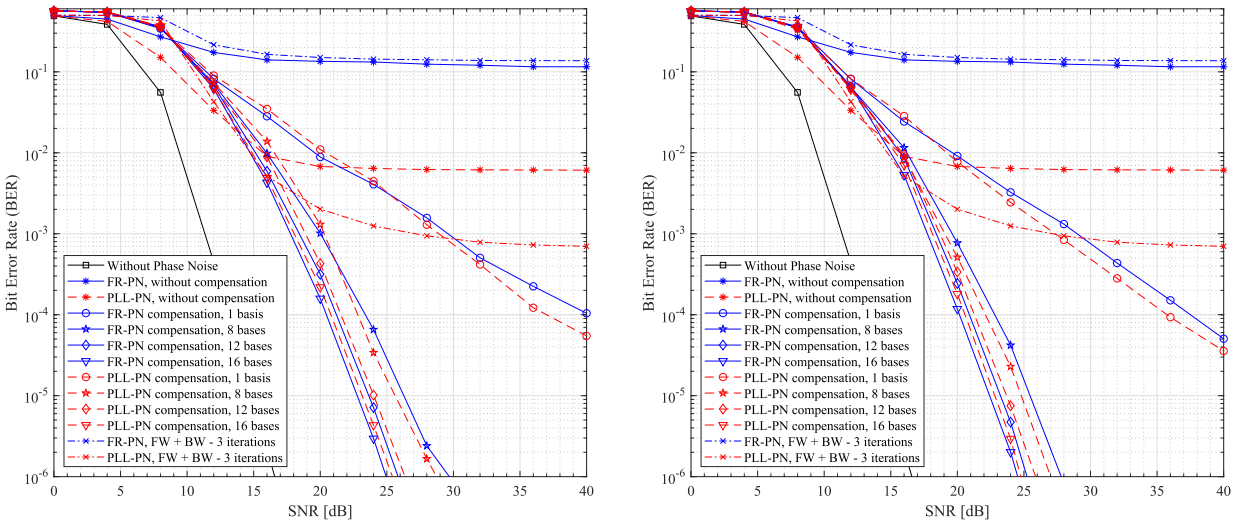


FIGURE 6. A comparison of squared errors between $(\mathbf{y}_1, \bar{\mathbf{y}}_1)$ and $(\tilde{\mathbf{y}}_1, \bar{\mathbf{y}}_1)$ over 1000 preamble frames at SNR = 15 dB.

this, define $\bar{\mathbf{y}}_1 = \mathbf{P}_1^H \mathbf{y}_1$ and $\tilde{\mathbf{y}}_1 = \hat{\mathbf{P}}_1^H \mathbf{y}_1$, where $\mathbf{P}_1$ and $\hat{\mathbf{P}}_1$ are the true PN and the estimated PN obtained from (26), respectively. This means that $\bar{\mathbf{y}}_1$ is the received signal without any effect of PN, whereas $\tilde{\mathbf{y}}_1$ is the received signal that experiences PN impact and PN compensation. From (18), without PN, the estimated channel is $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(1)} = \mathbf{C}\bar{\mathbf{y}}_1$ where $\mathbf{C} = N\breve{\mathbf{F}}\boldsymbol{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H\left(N\mathbf{S}_1\breve{\mathbf{F}}\boldsymbol{\Sigma}\breve{\mathbf{F}}^H\mathbf{S}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F}$. In the presence of PN but without PN compensation, the estimated channel is obtained as $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(2)} = \mathbf{C}\mathbf{y}_1$. However, if the first stage is deployed, i.e., the estimated PN in (26) is substituted into (18), the estimated channel is $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(3)} = \mathbf{C}\tilde{\mathbf{y}}_1$. Obviously, whether channel estimate $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(2)}$ or $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(3)}$ is closer to $\hat{\mathbf{h}}_{f,\text{MMSE}}^{(1)}$ depends on whether $\mathbf{y}_1$ or $\tilde{\mathbf{y}}_1$ is closer



(a) Basis vectors are from DFT matrix

(b) Basis vectors are from the PN covariance matrix

FIGURE 7. BER performance of the two-stage algorithm over EPA channel when $K = 128$, $M = 6$ and Martin shaping filter is deployed. The pilot ratio is 1/8.

to $\bar{\mathbf{y}}_1$ in the mean squared error sense. To see this, Fig. 6 displays squared errors between $(\mathbf{y}_1, \bar{\mathbf{y}}_1)$, denoted as δ_1, and $(\tilde{\mathbf{y}}_1, \bar{\mathbf{y}}_1)$, denoted as δ_2, over 1000 preamble frames and for SNR = 15 dB. It can be seen that there are cases where δ_1 is lower than δ_2, which means that the estimated PN in the first stage sometimes does not help to improve the received signal quality after performing PN compensation and thus does not improve channel estimation. However, the average value of δ_2, indicated in the figure as Δ_2, is considerably lower than the average value of δ_1, indicated as Δ_1. Therefore, on the average, the estimated PN still results in a lower MSE of the estimated channel as observed in Fig. 5.

C. STAGE 2: DATA DETECTION AND PHASE NOISE COMPENSATION

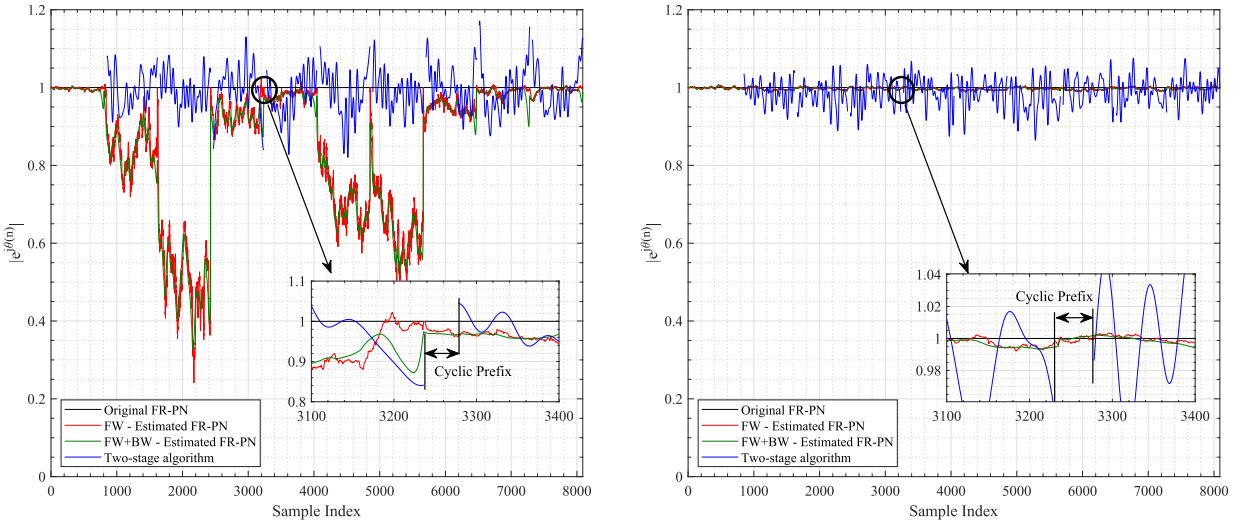
To compensate for the PN and detect data, pilots are inserted into transmission frames. Two pilot patterns are considered in this paper. For the first pattern, 15 pilot symbols are inserted into each time slot of data transmission blocks at positions [7, 15, 23, 31, 39, 47, 55, 63, 71, 79, 87, 95, 103, 111, 119], which corresponds to the pilot ratio of approximately 1/8. The second pattern has a pilot ratio of 1/16 in which 8 pilots are added at positions [7, 23, 39, 55, 71, 87, 103, 119].

BER performance of CFBMC-OQAM is first illustrated in Fig. 7 corresponding to when EPA channel is considered and the pilot ratio is 1/8. In this figure, performance of the case labeled as "Without Phase Noise" where PN is absent serves as the lower bound. There are two sub-figures in Fig. 7. The sub-figure on the left is for the case that F basis vectors are taken from N columns of the DFT matrix, while the sub-figure on the right is for the case when F basis vectors are taken from N eigenvectors of the PN covariance matrix $\mathbf{R}$ corresponding to the largest eigenvalues.

For comparison, Fig. 7 also includes performance obtained after 3 iterations of the iterative PN compensation algorithm proposed in [35]. In each iteration of such an algorithm, a forward LMS procedure is first applied to estimate the PN which is then smoothened by a backward procedure. As such, the algorithm proposed in [35] is referred to as "FW+BW" in Fig. 7. As suggested in [35], the estimation step sizes for the forward and backward procedures are selected as $\mu_\phi = 0.1$ and $\lambda_\phi = 1 - \mu_\phi$, respectively.

First, without PN compensation, the impact of FR-PN is much more severe than that from PLL-PN as expected. It can be seen from Fig. 7 that PN compensation performance is significantly improved when the number of bases increases. Specifically, for FR-PN distortion, at BER $= 10^{-5}$, performance gap between the ideal case, i.e., "Without Phase Noise", and "FR-PN compensation" reduces from 11.4 dB to 7.9 dB when 8 and 16 DFT-matrix-based basis vectors are considered, respectively. Similarly, for PLL-PN distortion, approximately 2.5 dB performance gain is obtained when the number of DFT-based bases increases from 8 to 16 vectors. About 8.3 dB performance gap is observed between the ideal case and "PLL-PN compensation, 16 bases". A similar trend can be seen in Fig. 7b when the basis vectors are taken from the PN covariance matrix. However, performance of the proposed algorithm using the PN covariance matrix tends to be better than using DFT matrix, especially when a low number of basis vectors is considered. It can also be seen that increasing the number of bases from 12 to 16 does not improve much the performance of the proposed two-stage PN compensation algorithm. In particular, when the bases are obtained from the PN covariance matrices, only approximately 0.8 dB gain is observed at BER $= 10^{-5}$ for both PN models.

Fig. 7 shows that applying the algorithm proposed in [35] does not result in good performance. To explain this, Fig. 8 (which contains two sub-figures) illustrate magnitudes of the estimated FR-PN in the forms of exponential function



(a) $c_v = 5 \times 10^{-18}$ (b) $c_v = 1 \times 10^{-19}$

FIGURE 8. A comparison in the PN estimation between the proposed algorithm in this paper and the iterative algorithm in [35]. For the proposed algorithm in this paper, 16 bases are deployed which are taken from DFT matrix. The pilot ratio is 1/8 and SNR = 25 dB.

obtained from the proposed algorithm in the second stage in this paper (labeled as "Two-stage algorithm"), the algorithm in [35] when only the FW procedure is considered (labeled as "FW - Estimated FR-PN"), and the algorithm in [35] when both FW and BW procedures are considered (labeled as "FW+BW - Estimated FR-PN"). The sub-figure on the left is obtained when the quality factor $c_v = 5 \times 10^{-18}$, while the right sub-figure is for $c_v = 1 \times 10^{-19}$. The first quality factor corresponds to FR-PN having a 3-dB bandwidth $\Delta f_{3\text{dB}}$ of over one thousand hertz, whereas the second quality factor corresponds to $\Delta f_{3\text{dB}}$ of dozens hertz.[4]

As can be seen from Fig. 8b, when the PN is small, the algorithm in [35] outperforms the algorithm proposed in this paper in tracking the PN. However, when the PN is large, significant fluctuations are observed for the estimated PN obtained from [35] in some specific positions as highlighted in Fig. 8a. It is worth noting that in [35], the PN compensation algorithm is performed for every transmission block and the PN corresponding to a symbol interval is estimated based on the estimated PN obtained for the previous symbol interval and the decoded signal. Thus, having decoding errors at the output of the channel decoder results in poor estimates of the PN. These errors consequently affect the PN estimation for the next symbol intervals. At the end, the errors are spread all over the whole transmission block. This error propagation phenomenon often happens when the decoding output is reused in an iterative algorithm. Consistent with the results illustrated in [35], performance of the iterative algorithm in terms of PER is still high due to the in-block error spreading. Another disadvantage of the iterative algorithm in [35] is the long latency.

[4]In FR-PN model, the 3-dB bandwidth $\Delta f_{3\text{dB}}$ is related to the quality factor c_v as $\Delta f_{3\text{dB}} \sim 2\pi c_v f_c^2$ [25], [34].

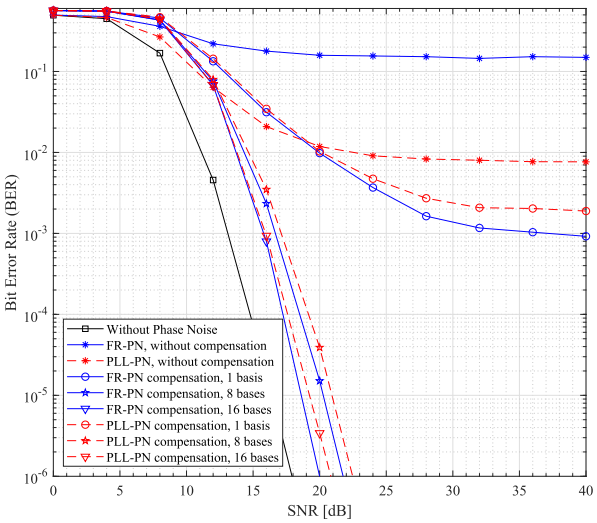
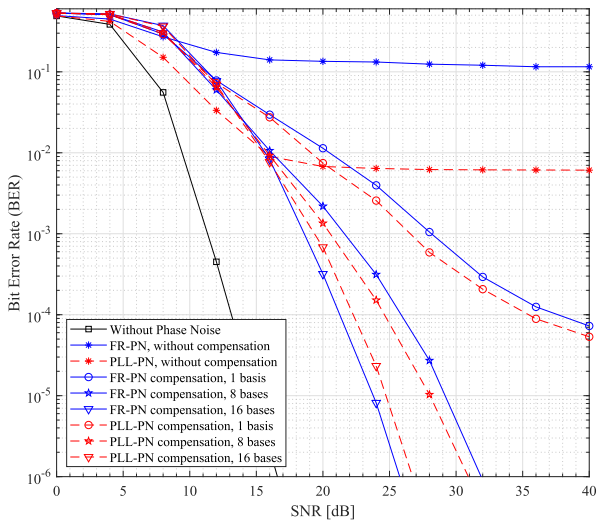


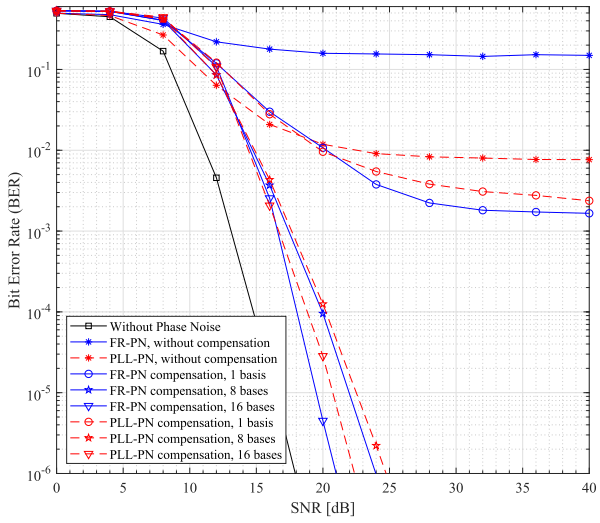
FIGURE 9. BER performance of the two-stage algorithm over EVA channel when $K = 128$, $M = 6$ and Martin shaping filter is deployed. Basis vectors are from the PN covariance matrix. The pilot ratio is 1/8.

Fig. 9 illustrates performance of the proposed algorithm when the EVA channel is considered. For this figure, the basis vectors are generated from the PN covariance matrix and the pilot ratio is 1/8. It can be seen that increasing the number of bases improves the performance and the improvement is better than when the EPA channel is considered. At BER $= 10^{-5}$, only about 2.5 dB performance gap is observed when 16 basis vectors are applied for two PN types compared to the ideal case.

Performance of the proposed algorithm for both EPA and EVA channels when the pilot ratio is 1/16 is presented in Fig. 10. In this figure, all bases are from the PN covariance matrix. As expected, decreasing the pilot ratio from 1/8 to



(a) EPA channel



(b) EVA channel

FIGURE 10. BER performance of the two-stage algorithm when $K = 128$, $M = 6$ and Martin shaping filter is deployed. Basis vectors are from the PN covariance matrix. The pilot ratio is 1/16.

1/16 degrades performance of the proposed algorithm. For the EPA channel, the performance gaps between the proposed algorithm with 16 basis vectors and the ideal case without PN are about 8.9 dB and 9.9 dB for FR-PN and PLL-PN, respectively, at BER $= 10^{-5}$ as seen in Fig. 10a. For the EVA channel in Fig. 10b, these gaps are 3.1 dB and 4.3 dB for FR-PN and PLL-PN, respectively.

It can be seen from Figs. 7, 9 and 10 that the BER performance of the proposed two-stage algorithm is degraded in the low SNR range compared to when no compensation algorithm is deployed. First, this phenomenon comes directly from the degradation of the estimated channel MSE in the first stage due to the under-determination problem discussed earlier. Furthermore, the proposed PN estimation in the second stage is essentially similar to the approximation problem in which the PN is approximated by a high-degree polynomial based on a set of equispaced interpolation points which are the pilot symbols in this paper. It will be shown later that this approximation results in a high fluctuation of the estimated PN in the low SNR range and thus degrades the compensation performance.

Finally, the PN estimation based on pilots in the second stage is evaluated in Fig. 11 when SNR $=$ 15 dB and SNR $=$ 35 dB. Specifically, the PN in the second transmission block is plotted in Fig. 11. Thus, the symbol index in Fig. 11 runs from 851 to 1618. It can be seen that while the PN distortion is estimated very coarsely when SNR $=$ 15 dB, the PN estimation is getting much better for both types of PN when SNR increases to 35 dB. From the results in Fig. 11, it appears that the PN is approximated by a high-degree curve. This is inline with the previous discussion on the similarity between the proposed PN estimation in the second stage and the approximation problem based on a set of equispaced interpolation points.

D. SPECTRAL EFFICIENCY

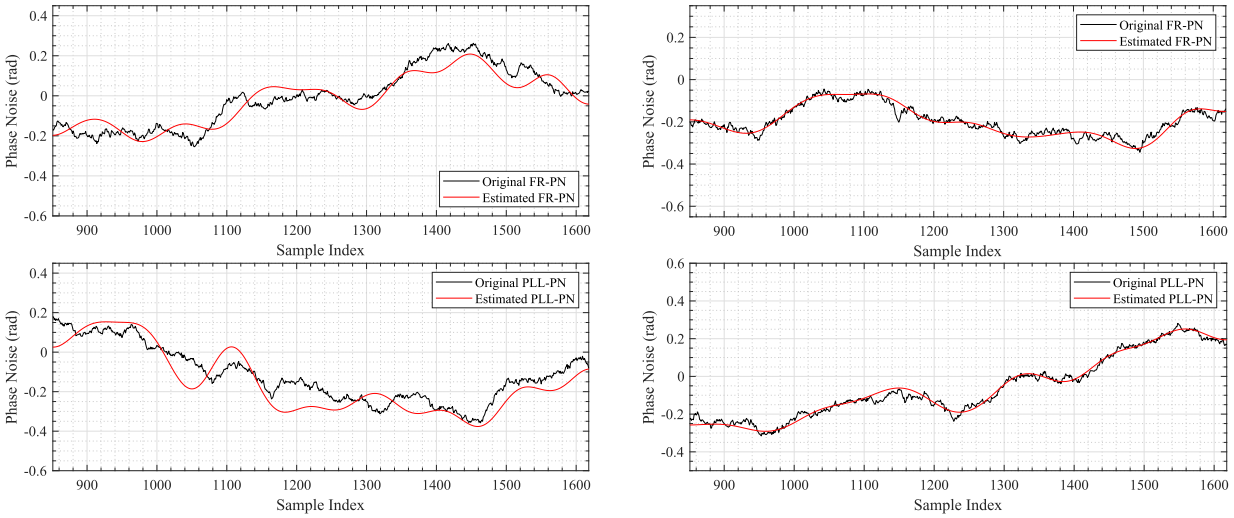
In general, the proposed algorithm trades bandwidth efficiency for better PN compensation. For the specific system parameters considered in the simulation with $K = 128$ subcarriers, $2 \times M = 12$ time slots, there are $N = K \times M = 768$ QAM symbols in each block, which also means that $N = 768$ QAM symbols are used in the preamble for the first stage. For the simulation results in Fig. 7 with a pilot ratio of 1/8, $N_p = 15 \times 6 = 90$ QAM pilot symbols are transmitted to estimate the PN in the second stage. Overall, the total number of QAM symbols used as preamble and pilots in the proposed two-stage algorithm is:

$$N_{\text{tot}}^{(1)} = N + (N_b - 1) \times N_p = 768 + (10 - 1) \times 90 = 1578. \tag{34}$$

In [35], before performing the iterative PN compensation, initial channel estimation is performed based on the transmission of preambles. Then a CPE compensation operation is performed for every remaining transmission block in a frame based on the transmission of pilots. It is presented in [35] that a two-symbol preamble, which is equivalent to two transmission blocks in this paper, is transmitted in each frame for initial channel estimation. Using the pilot ratio of 1/32 as in [35], $N_p = 1/32 \times 768 = 24$ pilots are needed for CPE compensation per each transmission block. Therefore, the total number of QAM symbols used as preamble and pilots in the algorithm in [35] is:

$$N_{\text{tot}}^{(2)} = 2 \times N + (N_b - 2) \times N_p = 1728. \tag{35}$$

It can be seen from (34) and (35) that in this particular case the algorithm proposed in this paper has a better overall spectral efficiency than the one in [35]. However, if the number of blocks per frame increases, the algorithm in [35] will be better than the proposed two-state algorithm in terms of the spectral efficiency.



(a) SNR = 15 dB (b) SNR = 35 dB

FIGURE 11. Estimated PN from the second stage when $K = 128$, $M = 6$, $F = 16$ and basis vectors are from the PN covariance matrix. The pilot ratio is 1/8.

VI. CONCLUSION

This paper has investigated the impact of PN on CFBMC-OQAM systems in the presence of imperfect channel estimation. First, it has been shown that there is a considerable performance gap between CFBMC-OQAM and OFDM under imperfect channel estimation. As such, a preamble design based on LS estimation has been proposed to reduce such performance gap. Simulation results indicated that applying the designed preamble results in a similar performance between CFBMC-OQAM and OFDM in terms of BER and MSE. In the presence of PN, the paper also proposed a two-stage algorithm to effectively estimate the channel frequency response and compensate for the PN impact. In particular, given the designed preamble, the channel frequency response and PN have been estimated in the first stage, in which the MMSE estimation method is employed to improve the channel estimation. In the second stage, the PN and data are estimated and detected based on the estimated channel obtained in the first stage and pilot symbols. Performance of the proposed two-stage PN compensation algorithm has been verified in a various number of practical simulation scenarios. Simulation results clearly shown that the two-stage algorithm effectively estimates the channel frequency response and compensates for the impact of PN.

APPENDIXES

A. APPENDIX 1

Based on the received signal in (16), the MMSE channel is $\hat{\mathbf{h}}_{f,\mathrm{MMSE}} = \hat{\mathbf{W}}\mathbf{y}_1$ where $\hat{\mathbf{W}}$ is a $N \times N$ weighting matrix that minimizes $\mathrm{E}\left[||\mathbf{h}_f - \hat{\mathbf{h}}_{f,\mathrm{MMSE}}||_2^2\right]$. Substituting $\hat{\mathbf{h}}_{f,\mathrm{MMSE}} = \hat{\mathbf{W}}\mathbf{y}_1$ into the orthogonality condition in (17) results in

$$\mathrm{E}\left[\left(\mathbf{h}_f - \hat{\mathbf{W}}\mathbf{y}_1\right)\mathbf{y}_1^H\right] = \mathbf{0} \Leftrightarrow \mathrm{E}\left[\mathbf{h}_f\mathbf{y}_1^H\right] - \hat{\mathbf{W}}\mathrm{E}\left[\mathbf{y}_1\mathbf{y}_1^H\right] = \mathbf{0} \quad (36)$$

Considering the first term in (36), one has

$$\begin{aligned}\mathrm{E}\left[\mathbf{h}_f\mathbf{y}_1^H\right] &\overset{(16)}{=} \mathrm{E}\left[\mathbf{h}_f\left(\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f + \mathbf{w}_1\right)^H\right] \\ &= \mathrm{E}\left[\mathbf{h}_f\mathbf{h}_f^H\right]\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H \\ &= N\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H \end{aligned} \quad (37)$$

The last equation in (37) is obtained based on the fact that $\mathbf{h}_f = \sqrt{N}\check{\mathbf{F}}\mathbf{h}$.

The second term in (36) can be derived as

$$\begin{aligned}&\hat{\mathbf{W}}\mathrm{E}\left[\mathbf{y}_1\mathbf{y}_1^H\right] \\ &\overset{(16)}{=} \hat{\mathbf{W}}\mathrm{E}\left[\left(\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f + \mathbf{w}_1\right)\left(\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f + \mathbf{w}_1\right)^H\right] \\ &= \hat{\mathbf{W}}\mathrm{E}\left[\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\mathbf{h}_f\mathbf{h}_f^H\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H + \mathbf{w}_1\mathbf{w}_1^H\right] \\ &= \hat{\mathbf{W}}\left(N\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right) \end{aligned} \quad (38)$$

From (36), (37), and (38), one obtains the weighting matrix $\hat{\mathbf{W}}$ as (39).

$$\begin{aligned}\hat{\mathbf{W}} &= N\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H\left(N\mathbf{P}_1\mathbf{F}^H\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\mathbf{F}\mathbf{P}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1} \\ &= N\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H\left(N\mathbf{S}_1\check{\mathbf{F}}\boldsymbol{\Sigma}\check{\mathbf{F}}^H\mathbf{S}_1^H + \sigma_{\mathbf{w}}^2\mathbf{I}\right)^{-1}\mathbf{F}\mathbf{P}_1^H \end{aligned} \quad (39)$$

The last equation in (39) is derived based on the fact that given two arbitrary matrices $\mathbf{R} \in \mathbb{C}^{N\times M}$ and $\mathbf{Q} \in \mathbb{C}^{M\times N}$, then the following identity is always true:

$$(\mathbf{I}_{N\times N} + \mathbf{R}\mathbf{Q})^{-1}\mathbf{R} = \mathbf{R}(\mathbf{I}_{M\times M} + \mathbf{Q}\mathbf{R})^{-1} \quad (40)$$

Therefore, the MMSE channel estimate is finally given as in (18).

B. APPENDIX 2

Define $\mathbf{z} = \mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t$. Then the optimization problem in (25) can be rewritten as

$$\begin{aligned} &\underset{\boldsymbol{\rho}}{\text{minimize}} \frac{1}{2}\left(\boldsymbol{\rho}^H\mathbf{z}+\mathbf{z}^H\boldsymbol{\rho}\right) \\ &\text{subject to } |\rho_i| = 1; \quad i = 0, 1, \cdots, N-1 \end{aligned} \tag{41}$$

Because of the constraint in (41), the solution, if exist, should be in the form of an exponential function. Define ε_i where $i = 0, 1, \cdots, N-1$, as Lagrange multipliers. Then the Lagrange function can be written as

$$\begin{aligned} C(\boldsymbol{\rho}, \boldsymbol{\varepsilon}) &= \frac{1}{2}\left(\boldsymbol{\rho}^H\mathbf{z}+\mathbf{z}^H\boldsymbol{\rho}\right) + \sum_{i=0}^{N-1}\varepsilon_i\left(||\rho_i||_2^2 - 1\right) \\ &= \frac{1}{2}\sum_{i=0}^{N-1}\left(\rho_i^* z_i + z_i^*\rho_i\right) + \sum_{i=0}^{N-1}\varepsilon_i\left(||\rho_i||_2^2 - 1\right) \\ &= \frac{1}{2}\sum_{i=0}^{N-1}\left(\left(\Re(\rho_i) - j\Im(\rho_i)\right)z_i + \right. \\ &\quad \left. z_i^*\left(\Re(\rho_i) + j\Im(\rho_i)\right)\right) \\ &\quad + \sum_{i=0}^{N-1}\varepsilon_i\left(\Re^2(\rho_i) + \Im^2(\rho_i) - 1\right) \end{aligned} \tag{42}$$

Taking the partial derivative of $C(\boldsymbol{\rho}, \boldsymbol{\varepsilon})$ with respect to $\Re(\rho_i)$ and $\Im(\rho_i)$ results in

$$\frac{\partial C(\boldsymbol{\rho}, \boldsymbol{\varepsilon})}{\partial \Re(\rho_i)} = \frac{1}{2}\left(z_i + z_i^*\right) + 2\varepsilon_i\Re(\rho_i) = 0 \tag{43a}$$

$$\frac{\partial C(\boldsymbol{\rho}, \boldsymbol{\varepsilon})}{\partial \Im(\rho_i)} = \frac{1}{2}j\left(-z_i + z_i^*\right) + 2\varepsilon_i\Im(\rho_i) = 0 \tag{43b}$$

which yields

$$\Re(\rho_i) = -\frac{\Re(z_i)}{2\varepsilon_i} \tag{44a}$$

$$\Im(\rho_i) = -\frac{\Im(z_i)}{2\varepsilon_i} \tag{44b}$$

where $i = 0, 1, \cdots, N-1$. Thus

$$\arg(\rho_i) = \frac{\Im(\rho_i)}{\Re(\rho_i)} = \frac{\Im(z_i)}{\Re(z_i)} = \arg(z_i) \tag{45}$$

Hence, the solution for the optimization problem in (25) is

$$\boldsymbol{\rho}_{t+1} = \mathrm{e}^{j\arg(\mathbf{z})} = \mathrm{e}^{j\arg(\mathbf{Y}_1^H(\mathbf{U}-\mathbf{V})\mathbf{Y}_1\boldsymbol{\rho}_t)} \tag{46}$$

REFERENCES

[1] B. Farhang-Boroujeny, "OFDM versus filter bank multicarrier," *IEEE Signal Process. Mag.*, vol. 28, no. 3, pp. 92–112, May 2011.
[2] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and its wireless applications: A survey," *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1673–1694, May 2009.
[3] M. Morelli, C.-C.-J. Kuo, and M.-O. Pun, "Synchronization techniques for orthogonal frequency division multiple access (OFDMA): A tutorial review," *Proc. IEEE*, vol. 95, no. 7, pp. 1394–1427, Jul. 2007.
[4] K. Vasudevan, "Coherent detection of turbo coded OFDM signals transmitted through frequency selective Rayleigh fading channels," in *Proc. IEEE Int. Conf. Signal Process., Comput. Control (ISPCC)*, Sep. 2013, pp. 1–6.
[5] E. P. Simon, L. Ros, H. Hijazi, and M. Ghogho, "Joint carrier frequency offset and channel estimation for OFDM systems via the EM algorithm in the presence of very high mobility," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 754–765, Feb. 2012.
[6] O. Weikert, "Joint estimation of carrier and sampling frequency offset, phase noise, IQ offset and MIMO channel for LTE advanced UL MIMO," in *Proc. IEEE 14th Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2013, pp. 520–524.
[7] K. Vasudevan, "Coherent detection of turbo-coded OFDM signals transmitted through frequency selective Rayleigh fading channels with receiver diversity and increased throughput," *Wireless Pers. Commun.*, vol. 82, no. 3, pp. 1623–1642, Jun. 2015.
[8] B. Farhang-Boroujeny, "Filter bank multicarrier modulation: A waveform candidate for 5G and beyond," *Adv. Electr. Eng.*, vol. 2014, Dec. 2014, Art. no. 482805.
[9] K. W. Martin, "Small side-lobe filter design for multitone data-communication applications," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 45, no. 8, pp. 1155–1161, Aug. 1998.
[10] K. Vasudevan, *Digital Communications and Signal Processing*. Boca Raton, FL, USA: CRC Press, 2010. [Online]. Available: https://dl.acm.org/doi/book/10.5555/1571785
[11] B. Farhang-Boroujeny, *Signal Processing Techniques for Software Radios*. Morrisville, NC, USA: Lulu Publishing House, 2011.
[12] A. Rezazadehreyhani, A. Farhang, and B. Farhang-Boroujeny, "Circularly pulse-shaped waveforms for 5G: Options and comparisons," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.
[13] G. Fettweis, M. Krondorf, and S. Bittner, "GFDM–generalized frequency division multiplexing," in *Proc. VTC Spring - IEEE 69th Veh. Technol. Conf.*, Apr. 2009, pp. 1–4.
[14] N. Michailow, M. Matthe, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 9, pp. 3045–3061, Sep. 2014.
[15] A. Rezazadeh Reyhani and B. Farhang-Boroujeny, "An analytical study of circularly pulse-shaped FBMC-OQAM waveforms," *IEEE Signal Process. Lett.*, vol. 24, no. 10, pp. 1503–1506, Oct. 2017.
[16] S. Adireddy, L. Tong, and H. Viswanathan, "Optimal placement of training for frequency-selective block-fading channels," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2338–2353, Aug. 2002.
[17] W. Zhang, X.-G. Xia, and P. C. Ching, "Optimal training and pilot pattern design for OFDM systems in Rayleigh fading," *IEEE Trans. Broadcast.*, vol. 52, no. 4, pp. 505–514, Dec. 2006.
[18] D. Katselis, E. Kofidis, A. Rontogiannis, and S. Theodoridis, "Preamble-based channel estimation for CP-OFDM and OFDM/OQAM systems: A comparative study," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2911–2916, May 2010.
[19] E. Kofidis, "Preamble-based estimation of highly frequency selective channels in FBMC/OQAM systems," *IEEE Trans. Signal Process.*, vol. 65, no. 7, pp. 1855–1868, Apr. 2017.
[20] J.-M. Choi, Y. Oh, H. Lee, and J.-S. Seo, "Pilot-aided channel estimation utilizing intrinsic interference for FBMC/OQAM systems," *IEEE Trans. Broadcast.*, vol. 63, no. 4, pp. 644–655, Dec. 2017.
[21] P. Singh, H. B. Mishra, A. K. Jagannatham, and K. Vasudevan, "Semi-blind, training, and data-aided channel estimation schemes for MIMO-FBMC-OQAM systems," *IEEE Trans. Signal Process.*, vol. 67, no. 18, pp. 4668–4682, Sep. 2019.
[22] P. Singh, H. B. Mishra, A. K. Jagannatham, K. Vasudevan, and L. Hanzo, "Uplink sum-rate and power scaling laws for multi-user massive MIMO-FBMC systems," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 161–176, Jan. 2020.
[23] S. Ehsanfar, M. Matthe, D. Zhang, and G. Fettweis, "Theoretical analysis and CRLB evaluation for pilot-aided channel estimation in GFDM," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.
[24] S. Ehsanfar, M. Matthe, D. Zhang, and G. Fettweis, "Interference-free pilots insertion for MIMO-GFDM channel estimation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–6.
[25] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1607–1616, Aug. 2007.
[26] S. Wu, P. Liu, and Y. Bar-Ness, "Phase noise estimation and mitigation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3616–3625, Dec. 2006.

[27] R. Corvaja and A. G. Armada, "Joint channel and phase noise compensation for OFDM in fast-fading multipath applications," *IEEE Trans. Veh. Technol.*, vol. 58, no. 2, pp. 636–643, Feb. 2009.

[28] R. A. Casas, S. L. Biracree, and A. E. Youtz, "Time domain phase noise correction for OFDM signals," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 230–236, Sep. 2002.

[29] A. Leshem and M. Yemini, "Phase noise compensation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 65, no. 21, pp. 5675–5686, Nov. 2017.

[30] S. Wu and Y. Bar-Ness, "A phase noise suppression algorithm for OFDM-based WLANs," *IEEE Commun. Lett.*, vol. 6, no. 12, pp. 535–537, Dec. 2002.

[31] P. Rabiei, W. Namgoong, and N. Al-Dhahir, "A non-iterative technique for phase noise ICI mitigation in packet-based OFDM systems," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5945–5950, Nov. 2010.

[32] P. Mathecken, T. Riihonen, S. Werner, and R. Wichman, "Phase noise estimation in OFDM: Utilizing its associated spectral geometry," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 1999–2012, Apr. 2016.

[33] Q. Zou, A. Tarighat, and A. H. Sayed, "Compensation of phase noise in OFDM wireless systems," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5407–5424, Nov. 2007.

[34] Z. Wang, P. Babu, and D. P. Palomar, "Effective low-complexity optimization methods for joint phase noise and channel estimation in OFDM," *IEEE Trans. Signal Process.*, vol. 65, no. 12, pp. 3247–3260, Jun. 2017.

[35] S. Suyama, H. Suzuki, K. Fukawa, and J. Izumi, "Iterative receiver employing phase noise compensation and channel estimation for millimeter-wave OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 8, pp. 1358–1366, Oct. 2009.

[36] S. Suyama, Y. Hashimoto, H. Suzuki, and K. Fukawa, "60 GHz OFDM experimental system employing decision-directed phase noise compensation," in *Proc. IEEE Radio Wireless Symp.*, Jan. 2012, pp. 191–194.

[37] D. Shin, S. Suyama, H. Suzuki, and K. Fukawa, "10 gbps millimeter-wave OFDM experimental system with iterative phase noise compensation," in *Proc. IEEE Radio Wireless Symp.*, Jan. 2013, pp. 184–186.

[38] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[39] V. Moles-Cases, A. A. Zaidi, X. Chen, T. J. Oechtering, and R. Baldemair, "A comparison of OFDM, QAM-FBMC, and OQAM-FBMC waveforms subject to phase noise," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.

[40] A. Kakkavas, M. Castañeda, J. Luo, T. Laas, W. Xu, and J. A. Nossek, "FBMC-OQAM with phase noise: Achievable performance and compensation," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2017, pp. 1–5.

[41] B. Lim and Y. C. Ko, "SIR Analysis of OFDM and GFDM Waveforms With Timing Offset, CFO, and Phase Noise," *IEEE Trans. Wireless Commun.*, vol. 16, no. 10, pp. 6979–6990, Oct. 2017.

[42] L. D. Le and H. H. Nguyen, "Impacts of phase noise on CFBMC-OQAM," in *Proc. IEEE 88th Veh. Technol. Conf. (VTC-Fall)*, Aug. 2018, pp. 1–6.

[43] P. Singh, E. Sharma, K. Vasudevan, and R. Budhiraja, "CFO and channel estimation for frequency selective MIMO-FBMC/OQAM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 844–847, Oct. 2018.

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

[45] A. Mehrotra, "Noise analysis of phase-locked loops," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 9, pp. 1309–1316, Sep. 2002.

[46] J.-J. van de Beek, O. Edfors, M. Sandell, S. K. Wilson, and P. O. Borjesson, "On channel estimation in OFDM systems," in *Proc. IEEE 45th Veh. Technol. Conf.*, vol. 2, Jul. 1995, pp. 815–819.

[47] Z. Cheng and D. Dahlhaus, "Time versus frequency domain channel estimation for OFDM systems with antenna arrays," in *Proc. 6th Int. Conf. Signal Process.*, vol. 2, Aug. 2002, pp. 1340–1343.

[48] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[49] S. Ohno, "Preamble and pilot symbol design for channel estimation in OFDM," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, vol. 3, Apr. 2007, pp. III-281–III-284.

[50] D. Katselis, "Some preamble design aspects in CP-OFDM systems," *IEEE Commun. Lett.*, vol. 16, no. 3, pp. 356–359, Mar. 2012.

[51] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation, Version 12.4.0*, document Technical Specification (TS) 36.211, 3rd Generation Partnership Project (3GPP), Feb. 2015.

[52] *IEEE Standard for Information Technology—Local and Metropolitan Area Networks—Specific Requirements—Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput*, IEEE Standard 802.11n-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, and IEEE Std 802.11w-2009) Oct. 2009, pp. 1–565.

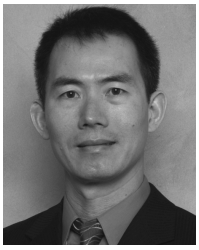
[53] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications With MATLAB*. Hoboken, NJ, USA: Wiley, 2010.

[54] *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) Radio Transmission and Reception, 3rd Generation Partnership Project (3GPP), Version 14.3.0.*, document Technical Specification (TS) 36.101, Apr. 2017.

[55] L. Zhang, P. Xiao, A. Zafar, A. U. Quddus, and R. Tafazolli, "FBMC System: An Insight Into Doubly Dispersive Channel Impact," *IEEE Trans. Veh. Technol.*, vol. 66, no. 5, pp. 3942–3956, May 2017.

[56] J. Zhang, M. Zhao, L. Zhang, J. Zhong, and T. Yu, "Circular convolution filter bank multicarrier (FBMC) system with index modulation," in *Proc. IEEE 86th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2017, pp. 1–5.

LONG D. LE received the B.Eng. degree in electronics and telecommunication from the Hanoi University of Technology (HUT), Hanoi, Vietnam, in 2011, and the M.Sc. degree in electrical and information the from Seoul National University of Science and Technology, Seoul, South Korea, in 2015. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Saskatchewan, Canada. His research interests include wireless communications, signal processing, full-duplex communications, and new multicarrier communications.



HA H. NGUYEN (Senior Member, IEEE) received the B.Eng. degree from the Hanoi University of Technology (HUT), Hanoi, Vietnam, in 1995, the M.Eng. degree from the Asian Institute of Technology (AIT), Bangkok, Thailand, in 1997, and the Ph.D. degree from the University of Manitoba, Winnipeg, MB, Canada, in 2001, all in electrical engineering. He joined the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK, Canada, in 2001, where he became a Full Professor, in 2007. He is currently the NSERC/Cisco Industrial Research Chair of low-power wireless access for sensor networks. He has coauthored, with E. Shwedyk, the textbook *A First Course in Digital Communications* (Cambridge University Press). His research interests include communication theory, wireless communications, and statistical signal processing. He is a Fellow of the Engineering Institute of Canada (EIC) and a Registered Member of the Association of Professional Engineers and Geoscientists of Saskatchewan (APEGS). He was the Co-Chair of the Multiple Antenna Systems and Space-Time Processing Track of the IEEE Vehicular Technology Conferences (in Fall 2010, Ottawa, ON, Canada, and in Fall 2012, Quebec, QC, Canada), the Lead Co-Chair of the Wireless Access Track of the IEEE Vehicular Technology Conferences (in Fall 2014, Vancouver, BC, Canada), the Lead Co-Chair of the Multiple Antenna Systems and Cooperative Communications Track of the IEEE Vehicular Technology Conference (in Fall 2016, Montreal, QC, Canada), and the Technical Program Co-Chair of the Canadian Workshop on Information Theory (St. John's, NL, Canada, in 2015). He was an Associate Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the IEEE WIRELESS COMMUNICATIONS LETTERS, from 2007 to 2011 and from 2011 to 2016, respectively. He also serves as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.