WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2016; **16**:1018–1034
Published online 3 March 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/wcm.2582

RESEARCH ARTICLE

Low complexity estimation of carrier and sampling frequency offsets in burst-mode OFDM systems

Yonathan Murin* and Ron Dabora

Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, 8410501, Israel

ABSTRACT

Residual carrier frequency offset (CFO) and sampling frequency offset (SFO) in orthogonal frequency-division multiplexing (OFDM) transmission result in loss of orthogonality between the subcarriers, which severely degrades system performance. This makes effective CFO and SFO estimation methods essential for such systems. In this work, we address the joint estimation of CFO and SFO, focusing on burst-mode OFDM communications scenarios.

We study communications under slowly varying channels, and consider three cases of knowledge of the channel impulse response (CIR): full knowledge, no knowledge, and partial knowledge of the CIR. By partial knowledge, we refer to knowing only either the CIR magnitudes or the CIR phases. It is known that obtaining the exact joint maximum-likelihood estimate (MLE) of the CFO and the SFO requires a two-dimensional search. Here, we present a new estimation method which uses the Taylor expansion of the MLE cost function, combined with the best linear unbiased estimator, to obtain a method which does not require such a search. The computational complexity of the new method is evaluated. Numerical simulations demonstrate that the new method approaches the corresponding Cramér-Rao bound for a wide range of signal-to-noise ratios, and has superior performance compared to all other existing methods for approximating the solution for the joint MLE, while maintaining a low computational complexity. Copyright © 2015 John Wiley & Sons, Ltd.

KEYWORDS

orthogonal frequency-division multiplexing; carrier frequency offset estimation; sampling frequency offset estimation

***Correspondence**

Yonathan Murin, Department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, 8410501, Israel.
E-mail: moriny@ee.bgu.ac.il

1. INTRODUCTION

Orthogonal frequency-division multiplexing (OFDM) is widely used in many wireless communication systems and commercial standards, for example, digital video broadcasting (DVB-T) [1], IEEE 802.16 wireless metropolitan area network (WMAN) [2], and 3rd Generation Partnership Project (3GPP) long term evolution (LTE) [3]. This is due to the robustness and high spectral efficiency of OFDM-modulated signals in frequency-selective channels. However, a major drawback of OFDM is its sensitivity to synchronization errors [4,5]. Therefore, accurate synchronization is essential in order to benefit from the potential advantages of OFDM. A common practice in burst-mode OFDM systems is to first apply a coarse acquisition based on a training sequence located at the beginning of the OFDM frame, for example [6–10], and then apply fine synchronization based on pilots symbols which are multiplexed with the data, for example [11]. In this work, we study the joint fine estimation of carrier frequency offset (CFO) and sampling frequency offset (SFO) in burst-mode OFDM systems.

CFO and SFO are induced by a mismatch between oscillator frequencies at the transmitter and at the receiver. The effect of CFO and SFO was analyzed in [12], in which it was shown that they both destroy the orthogonality of the subcarriers, thereby resulting in inter-carrier interference (ICI). Furthermore, while CFO induces the same phase drift at all subcarriers, SFO causes a phase drift that grows linearly with the subcarrier index. We note that there is a fundamental difference between CFO and SFO estimation for scenarios in which the channel impulse response (CIR) is rapidly varying, and CFO and SFO estimation for constant or slowly varying CIR scenarios: In rapidly varying environments, the phase variations induced by the CFO and the SFO can be incorporated into the instantaneous CIR, but in slowly varying scenarios, the CIR has to be separately treated from the phase variations induced by the CFO and the SFO. In this work, we focus on slowly varying channels, that is, the CIR remains constant over several consecutive OFDM symbols. When the CIR

changes slowly, the CFO and SFO are typically estimated based on a set of dedicated pilots, see, e.g., [1–3]. The estimation problem can be further expanded to include the joint estimation of the CFO, the SFO and the CIR, see, for example, the works [13,14] and [15], and references therein.

The problem of estimating the CFO and SFO for *known* CIR received a lot of attention in recent years. In [16], Laourine *et al.* applied harmonic retrieval method to the estimation of the CFO and SFO. However, this approach requires half of the subcarriers to be left unmodulated, which substantially reduces the spectral efficiency of the transmitted signal. The uplink of multiuser OFDM systems was studied in [17], which introduced a maximum-likelihood (ML)-based method for jointly estimating the CFOs and SFOs of all users. The work [17] approximated the received signal model using the first order Taylor expansion and then obtained the joint ML estimator (MLE) for the CFOs and SFOs based on this approximated model. It should be noted that this approximation is accurate only for small values of CFO and SFO. Another ML-oriented method was presented by Oberli in [18] for the case of multiple-input multiple-output (MIMO) OFDM systems. Oberli suggested to estimate the CFO and SFO based on the differences between the phases of the received samples in the pilot subcarriers using the knowledge of the phases of the pilot symbols and channel coefficients. Lastly, the work [19] used the same approach as the one used in [18] for obtaining an estimation method, which is specifically tailored to the 3GPP LTE downlink standard.

Estimation of the CFO and SFO for *unknown* CIR was first studied by Speth *et al.* in [20]. The work [20] estimated the CFO and SFO via a normalized sum and difference of the phase differences between the pilot subcarriers computed over two consecutive OFDM symbols, which are set to contain identical pilot patterns. Because the estimators of [20] ignore other impairments, e.g., I/Q imbalance, it can be applied without modification in the presence of such impairments, see, e.g., [21] and [22]. However, the resulting performance are substantially suboptimal. Another estimation method for the case of unknown CIR was proposed in [23]. The work [23] again used the phase differences between two consecutive OFDM symbols that contain identical pilots patterns. Unlike [20], the work [23] used the least-squares (LS) approach for estimation. Similarly to [20] and [23], the work [24] also estimated the CFO and SFO using the phase differences between two consecutive OFDM symbols containing identical pilot patterns: first, a coarse estimation of the CFO was carried out, then a compensation in the time-domain was applied using the estimated CFO, and finally a fine joint estimation of the CFO and SFO was implemented via averaging the phase differences over the pilots subcarriers. Note that estimation based on phase differences, as implemented in [20,23] and [24], constitutes an ad hoc approach and does not carry optimality. ML-oriented estimation of the CFO and SFO without CIR knowledge was studied in [25]. The work [25] showed that the exact joint-ML estimation of the CFO and SFO, based on known pilots, requires a two-dimensional search, which is computationally too demanding to be implemented in practice. Therefore, [25] presented a method to decouple these two estimation problems, such that only a single-dimension search is required. Furthermore, unrelated to the ML-oriented scheme, [25] also established a closed-form reduced-complexity estimator for the CFO and SFO whose performance are close to the MLE performance, only when the signal-to-noise ratio (SNR) is medium-to-high. It thus follows that the estimation methods presented in [25] either use an exhaustive search or have a performance that is considerably worse than the MLE performance in the low SNR regime.

Another approach for the estimation of CFO and SFO for unknown CIR is to use an estimation scheme with a specifically designed training sequence. Along this line, we note the works [13,26] and [27], which considered joint estimation of the CFO and the SFO based on two long training symbols: The work [13] applied a recursive LS approach to jointly estimate the CIR, CFO and SFO; the work [26] proposed a joint CFO and SFO MLE in the frequency domain which uses a two-dimensional search; and, the work [27] followed the approach of [25], and proposed to replace the two-dimensional search of [26] with a decoupled estimator that consists of closed-form estimation of the CFO, followed by an approximate MLE for the SFO. This approximate MLE is obtained via a second order Taylor expansion of the MLE cost function for the SFO, computed from the initial joint MLE cost function by plugging the ML estimate of the CFO. Note that in contrast to the work [25] and to the approach presented in this manuscript, the estimators presented in the works [13,26] and [27] are applicable only to the setting of two identical consecutive training symbols. As training reduces the throughput, in the present work, we focus on estimation without the help of training.

Main contributions

In this work, we consider joint estimation of the CFO and the SFO for three CIR scenarios: full CIR knowledge, no CIR knowledge, and partial CIR knowledge. We introduce a new polynomial approximation for the ML cost function, which transforms the search for maximizing the cost function into the problem of finding the roots of a polynomial, for which there are efficient numerical solution methods. The proposed polynomial approximation is next used in the estimation of a linear combination of the CFO and SFO at each pilot subcarrier of the received OFDM symbols, and then the best linear unbiased estimator (BLUE) is applied to obtain the estimates of the CFO and SFO from the estimated linear combinations. Numerical simulations show that the performance of this scheme closely approaches the performance of the joint MLE of the CFO and SFO and are superior to all other existing schemes which aim at obtaining an approximate joint MLE solution. At the same time, the computational complexity of the new scheme is 60% of

the complexity associated with the schemes based on the decoupled MLE.

Another important result we present in this work is the fact that the optimal performance, that is, the Cramér-Rao bound (CRB), for the case of known channel magnitudes and unknown channel phases, is the same as for the case of completely unknown CIR, while for the case of known phases and unknown magnitudes, the CRB is the same as for the case of completely known CIR. This result indicates that in the context of estimating the CFO and SFO, knowledge of the CIR phases is significantly more beneficial compared with knowledge of the CIR gains, which agrees well with the intuition regarding this estimation problem.

We note here that in contrast to the estimation methods presented in [25], which either use an exhaustive search or have performance which are considerably worse than the MLE performance at the low SNR regime, the estimation method presented here approaches the MLE performance *without using an exhaustive search*, even in the low SNR regime. We also note that in contrast to the work [17], which used first order Taylor expansion to approximate the *received signal model*, in this study, we approximate the *ML cost function*, which results in an approximation that holds for larger values of CFO and SFO and has superior performance. This is clearly demonstrated in the simulation results. Finally, we note that opposed to the work [27], the estimation method developed in this work applies pilot-assisted joint estimation, and is *not restricted to a two consecutive OFDM symbols*. Moreover, while the work [27] used a second order Taylor expansion to approximate only the generalized ML cost function for the estimation of the SFO after plugging the estimation of the CFO as a function of the SFO (this can be done only since two consecutive symbols are considered), our approach is based on the *joint ML cost function for the CFO and SFO* for any number of OFDM symbols. Therefore, the estimation method developed in this work has superior performance compared with [27]. In fact, because [27] is restricted to only two symbols, it is inherently suboptimal to [25]. Thus, the comparison with [25] included in the simulation study also indicates the comparison between the present work and the approach of [27].

The rest of this paper is organized as follows: In Section 2 we introduce the signal model and the problem definition. Section 3 reviews some preliminary results. In Section 4 we present the polynomial approximation technique and our new estimation method for all CIR scenarios. We present simulation results in Section 5, and conclude the paper in Section 6.

Notations: We denote vectors with boldface letters, e.g., $\mathbf{X}$, matrices with doublestroke letters, e.g., $\mathbb{A}$, and sets with calligraphic letters, e.g., $\mathcal{A}$. We use $\mathrm{E}\{\cdot\}$ to denote expectation and $\mathrm{Var}\{\cdot\}$ to denote variance. We use $\mathfrak{C}$, $\mathfrak{R}$ and $\mathfrak{N}$ to denote the sets of complex numbers, real numbers, and positive integers, respectively. Complex conjugate, transposition, and Hermitian transposition are denoted by $(\cdot)^*$, $[\cdot]^T$ and $[\cdot]^H$, respectively, while $\Re\{\cdot\}$ and $\Im\{\cdot\}$ denote the real and imaginary parts of a complex number.

Computational complexity: We define computational complexity of a method as the total number of floating point operations (flops), real as well as complex, needed to complete its required calculations, and denote it by $\mathscr{C}_{\text{method}}$. The complexity analysis in this work is presented for the case of known CIR. Analysis for the case of unknown CIR can be obtained by following similar arguments.

2. SIGNAL MODEL AND PROBLEM DEFINITION

We consider the same signal model as in [12,13,18,26] and [25]: an OFDM signal is generated via a discrete Fourier transform (DFT) whose size is N. Each OFDM symbol modulates N_D data symbols, where N_D is an even number [12,13,18,25,26]. The data symbols are denoted with $D_{l,k}$, where l is the index of the OFDM symbol and $-N_D/2 \leq k \leq N_D/2$, is the index of the subcarrier frequency. The symbols $D_{l,k}$ are independent and identically distributed (i.i.d.). For the l'th OFDM symbol, the modulated discrete-time sequence is obtained from $\{D_{l,k}\}$ by applying an inverse DFT (IDFT) of size N. Intersymbol interference (ISI) between subsequent OFDM symbols is avoided by adding a guard interval, referred to as cyclic prefix (CP), of length N_g samples to the IDFT output sequence. The resulting OFDM symbol, of length $N_s = N + N_g$, is input to the channel at a rate of $\frac{1}{T}$. Let $T_u \triangleq NT$, $T_g \triangleq N_g T$, and $T_s = T_u + T_g$. The transmitted complex baseband signal is given by [12, Equation (1)]:

$$s(t) = \frac{1}{\sqrt{T_u}} \sum_{l=-\infty}^{\infty} \sum_{k=-N_D/2, k\neq 0}^{N_D/2} D_{l,k}\psi_{l,k}(t), \qquad (1)$$

where the subcarriers pulses $\psi_{l,k}(t)$ are [12, Equation (2)]:

$$\psi_{l,k}(t) = \exp\left(j\frac{2\pi k \cdot (t - T_g - lT_s)}{T_u}\right) \cdot p(t - lT_s), \quad (2)$$

and $p(t) = 1$ for $0 \leq t < T_s$ and $p(t) = 0$ otherwise. In order to avoid difficulties due to DC offset, we set $D_{l,0} = 0$, see [25] and [26] .

Burst mode OFDM transmission takes place in frames, where each frame consists of $L + 1$ OFDM symbols transmitted consecutively. We consider transmission over slowly varying frequency-selective fading channels such that the channel remains approximately constant over an OFDM frame. We further assume correct OFDM symbol time synchronization, which implies that there is no interference between adjacent OFDM symbols. Due to oscillators instability, there are CFO and SFO at the receiver (with respect to the transmitter). In this work, we focus on estimating the residual CFO and SFO after coarse acquisition has been applied at the beginning of the reception of OFDM frame. Thus, relatively small SFO and CFO are present in the signal.

Let $h(t)$ denote the CIR and let Δf [Hz] denote the carrier frequency offset between the transmitter's carrier frequency and the receiver's local oscillator used for the down conversion of the received RF signal to baseband. Let $*$ denote the convolution operator. The resulting received, continuous-time baseband signal is given by [18, Equation (2)] $r(t) = e^{j2\pi\Delta ft} \cdot (s(t) * h(t)) + v(t)$, where $v(t)$ is a zero-mean white Gaussian noise process, such that $\mathrm{E}\{\upsilon(t_1)\upsilon(t_2)\} = \sigma_\upsilon^2\delta(t_1 - t_2)$, and σ_υ^2 is a finite variance. The received signal is sampled at intervals T' which results in an SFO which equals to $\delta = (T' - T)/T, \delta \in \mathfrak{R}_\delta$, where $\mathfrak{R}_\delta$ is the possible range of δ considered by the estimator. Let $\epsilon = \Delta f \cdot T$ denote the normalized CFO, where $\mathfrak{R}_\epsilon$ denotes its possible range considered by the estimator, that is, $\epsilon \in \mathfrak{R}_\epsilon$. The ranges $\mathfrak{R}_\epsilon$ and $\mathfrak{R}_\delta$ are assumed to be symmetric around zero and are determined by the estimation scheme and by the system requirements regarding the initial error value and the allowed accuracy in the system's clocks and oscillators. After sampling and CP removal, demodulation via an N-point DFT is applied. Let H_k denote the DFT of the sampled CIR at subcarrier k, and let $\mathrm{ICI}_{l,k}$ denote the ICI at the k'th subcarrier of the l'th OFDM symbol. Applying steps similar to the steps leading to [18, Equations (10)-(13)], to the sampled received baseband signal $r(t)$, and using the approximation $\epsilon + \delta\epsilon + \delta k \approx \epsilon + \delta k$, which holds due to the values of ϵ and δ, we obtain the approximated discrete-time received signal model

$$Z_{l,k} = e^{j\xi_l(\epsilon+k\delta)} \frac{\sin(\pi(\epsilon + k\delta))}{N \sin\left(\frac{\pi(\epsilon+k\delta)}{N}\right)} D_{l,k}H_k + \mathrm{ICI}_{l,k} + V_{l,k}, \tag{3}$$

where $\xi_l = \pi\frac{N-1+2N_l}{N}, N_l = lN_s + N_g$, and $V_{l,k} \sim \mathcal{CN}\left(0, \sigma_\upsilon^2\right)$ is an i.i.d, zero mean, complex Normal random process with variance σ_υ^2. Next, note that the attenuation of the subcarriers magnitudes can be approximated by unity when $\epsilon + \delta k \ll 1$, [12]. Furthermore, from [12, Equation (42)] we have that $\mathrm{E}\left\{\mathrm{ICI}_{l,k}^2\right\} \approx \frac{\epsilon^2\pi^2}{3}$. Therefore, for medium SNR conditions, e.g., SNR = 20dB, and for small enough values of ϵ and δ, e.g., $\epsilon = 10^{-2}$ and $\delta = 10^{-4}$ (these values are relevant to practical scenarios and were also used in [25, Subsection VI.A], see a detailed scenario parameters and their relationship with the values of ϵ and δ in Subsection 5.1), the ICI term is negligible compared with the additive Gaussian noise. On the other hand, when the SNR is high, the term $\mathrm{ICI}_{l,k}$ is non negligible and results in a noise floor. This is verified in Subsection 5.2 ,where the SNR is taken as high as 40 dB.

We study the estimation of the CFO and SFO based on a set $\mathcal{K}_p = \{k_1, k_2, \ldots, k_{K_p}\}$ of pilot tones, which is available at each payload OFDM symbol. Let $l = 0, 1, \ldots, L$ be the index of the OFDM symbol in the frame. As the channel remains approximately constant over a frame, the CIR can be estimated using a preamble, transmitted at the first OFDM symbol of the frame. The CFO and SFO are estimated at each frame, based on the pilot subcarriers in OFDM symbols $l = 1, 2, \ldots, L$. Thus, we arrive at the following simplified received signal model, for the subcarriers indexed in the set $\mathcal{K}_p$ [18,25]:

$$Z_{l,k_p} = e^{j\xi_l(\epsilon+k_p\delta)} \cdot D_{l,k_p}H_{k_p} + V_{l,k_p}. \tag{4}$$

The symbols used for modulating the pilot subcarriers are set to have a constant envelope $|D_{l,k_p}|^2 = \sigma_{d_p}^2 = 1$. Thus, the SNR over the pilot subcarriers is given by $\frac{|H_{k_p}|^2}{\sigma_\upsilon^2}$. As in [18, Section III], we consider the worst case situation in which the CFO ϵ and the SFO δ are independent of each other. Finally, since $H_{k_p} \in \mathfrak{C}$, we can write $H_{k_p} = |H_{k_p}|e^{j\phi_{k_p}}$. Known CIR refers to knowing both $|H_{k_p}|$ and ϕ_{k_p}, for all $k_p \in \mathcal{K}_p$,[†] while partial CIR refers to knowing only $|H_{k_p}|$'s or ϕ_{k_p}'s , as will be defined.

3. PRELIMINARIES

We begin with a short description of the BLUE. A detailed description of the BLUE can be found in [28, Ch. 6].

3.1. A brief overview of the BLUE

Consider a vector of N measurements denoted by $\mathbf{x} = x_0, x_1, \ldots, x_{N-1}$, whose probability density function depends on an unknown parameter vector $\boldsymbol{\theta}$ of size $p \times 1$. A linear estimator of $\boldsymbol{\theta}$ based on the measurements vector $\mathbf{x}$ has the general form $\hat{\boldsymbol{\theta}} = \mathbb{A}\mathbf{x}$, where $\mathbb{A}$ is a $p \times N$ matrix. The BLUE is defined to be the estimator that has the minimum variance among all the unbiased linear estimators. If the measurements $\mathbf{x}$ and the unknown parameter vector $\boldsymbol{\theta}$ follow a general linear relationship $\mathbf{x} = \mathbb{H}\boldsymbol{\theta} + \mathbf{w}$, where $\mathbb{H}$ is a known $N \times p$ matrix, and $\mathbf{w}$ is an $N \times 1$ zero-mean noise vector with covariance matrix $\mathbb{Q}$, then the BLUE of $\boldsymbol{\theta}$ is given by [28, Thm. 6.1]:

$$\hat{\boldsymbol{\theta}} = \left(\mathbb{H}^T\mathbb{Q}^{-1}\mathbb{H}\right)^{-1}\mathbb{H}^T\mathbb{Q}^{-1}\mathbf{x},$$

and the covariance matrix of $\hat{\boldsymbol{\theta}}$ is given by:

$$\mathrm{E}\left\{\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\theta}}^T\right\} = \left(\mathbb{H}^T\mathbb{Q}^{-1}\mathbb{H}\right)^{-1} \triangleq \mathbb{Q}_{\hat{\boldsymbol{\theta}}}.$$

Let $\boldsymbol{\theta} = [\epsilon, \delta]^T \in \mathfrak{R}_\epsilon \times \mathfrak{R}_\delta$, denote the unknown desired parameters vector. We next review the previously derived MLEs and CRBs for estimating $\boldsymbol{\theta}$ based on the model (4).

[†]Note that in slowly varying channels, the channel can be constant over several OFDM frames. In such a case, high accuracy CIR knowledge, corresponding to known CIR, can be obtained via estimating the CIR based on several boosted preamble OFDM symbols, see, for example, [2] and [29].

3.2. MLEs and CRB Expressions

As $V_{l,k}$ are complex Normal i.i.d random variables (RVs), it follows that the log-likelihood function for $\boldsymbol{\theta}$ is [18, Equation (20)], [25, Equation (8)]:

$$\Gamma(\boldsymbol{\theta}) = -LK_p \ln\left(\pi\sigma_{\upsilon}^2\right) - \frac{1}{\sigma_{\upsilon}^2}\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\left|Z_{l,k_p} - e^{j\xi_l(\epsilon+k_p\delta)}\cdot D_{l,k_p}H_{k_p}\right|^2. \quad (5)$$

We begin with the known CIR case. Letting $Y_{l,k_p} = Z_{l,k_p}D_{l,k_p}^*H_{k_p}^*$, the MLE for the case of known CIR, $\hat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(\mathrm{k})}$, is given by $\underset{\boldsymbol{\theta}\in\mathfrak{R}_\epsilon\times\mathfrak{R}_\delta}{\operatorname{argmax}}\{\Gamma(\boldsymbol{\theta})\}$:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(\mathrm{k})} = \underset{\boldsymbol{\theta}\in\mathfrak{R}_\epsilon\times\mathfrak{R}_\delta}{\operatorname{argmax}}\left\{\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\Re\left\{Y_{l,k_p}e^{-j\xi_l(\epsilon+k_p\delta)}\right\}\right\}. \quad (6)$$

Observe that the exact MLE requires a two-dimensional exhaustive search in ϵ and δ. Such a search entails a high computational complexity, thereby rendering the exact MLE impractical for most communication systems. This motivates the study of simplified and efficient estimation methods, which are the focus of this work.

As a baseline for the performance of the estimators derived in this work, we use the CRB for estimating ϵ and δ based on (4), calculated assuming that the CIR is known. Define $Q(q) \triangleq \sum_{k_p\in\mathcal{K}_p}k_p^{q-1}|H_{k_p}|^2, q = 1,2,3$, $Q^\star(j) = \frac{Q(j)}{Q(1)Q(3)-Q^2(2)}, j = 1,3$, and let $\mathrm{F}_{N,L} = \sum_{l=1}^{L}(N-1+2N_l)^2$. The CRBs for the estimates of ϵ and δ are derived in Appendix A. The corresponding expressions are:

$$\mathrm{CRB}_{\epsilon}^{(\mathrm{k})} = \frac{N^2\sigma_{\upsilon}^2Q^\star(3)}{2\pi^2\mathrm{F}_{N,L}},\ \mathrm{CRB}_{\delta}^{(\mathrm{k})} = \frac{N^2\sigma_{\upsilon}^2Q^\star(1)}{2\pi^2\mathrm{F}_{N,L}}. \quad (7)$$

Note that as the model (4) neglects the ICI, the CRBs (7) do not account for ICI. Therefore, these bounds are not tight in the high SNR regime.

For the unknown CIR case, the MLE and the CRBs were obtained in [25]. Here, we briefly review these results. Let $X_{l,k_p} = Z_{l,k_p}D_{l,k_p}^*$ and $R_{l,k_p} = \sum_{n=l+1}^{L}X_{n-l,k_p}^*X_{n,k_p}, l = 1,2,\ldots,L-1$. The MLE for $\boldsymbol{\theta}$ is given by [25, Equation (14)]:

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}}^{(\mathrm{uk})} = \underset{\boldsymbol{\theta}\in\mathfrak{R}_\epsilon\times\mathfrak{R}_\delta}{\operatorname{argmax}}\left\{\sum_{l=1}^{L-1}\sum_{k_p\in\mathcal{K}_p}\Re\left\{R_{l,k_p}e^{-j\pi\frac{2lN_s}{N}(\epsilon+k_p\delta)}\right\}\right\}. \quad (8)$$

Note that (8) and (6) have the same form with Y_{l,k_p} in (6), replaced by R_{l,k_p} in (8). The MLE (6) uses its knowledge of the CIR for coherently canceling the phase of the channel coefficients. The MLE of (8), on the other hand, has no CIR knowledge. Therefore, as the CIR is constant over L consecutive OFDM symbols, the effect of the channel phase is canceled by using the phase differences for the same subcarrier index via R_{l,k_p}.

Next, define $\mathrm{J}_{N,L} = \mathrm{F}_{N,L} - \frac{1}{L}\left(\sum_{l=1}^{L}(N-1+2N_l)\right)^2$ The CRB for the estimation of $\boldsymbol{\theta}$ based on (4), assuming the CIR is unknown, is given by [25, Equations (12)–(13)]:

$$\mathrm{CRB}_{\epsilon}^{(\mathrm{uk})} = \frac{N^2\sigma_{\upsilon}^2Q^\star(3)}{2\pi^2\mathrm{J}_{N,L}},\ \mathrm{CRB}_{\delta}^{(\mathrm{uk})} = \frac{N^2\sigma_{\upsilon}^2Q^\star(1)}{2\pi^2\mathrm{J}_{N,L}}. \quad (9)$$

Similarly to the known CIR case, the CRBs (9) are not tight in the high SNR regime.

3.3. Decoupled MLEs

The MLE (8) can be simplified by decoupling the estimations of the CFO and of the SFO, following the procedure presented in [25, Subsection IV.A]. The main idea of the decoupled ML (DML) estimator is to expand the CFO parameter, ϵ, into a vector of parameters $\boldsymbol{\alpha}$, whose estimations can be decoupled from the estimation of the SFO δ. Define $\alpha_l = 2lN_s\epsilon$ and consider the vector $\boldsymbol{\alpha} = [\alpha_1,\alpha_2,\ldots,\alpha_L]$. The summands in (8) can be expressed using $\boldsymbol{\alpha}$ as:

$$\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\Re\left\{R_{l,k_p}e^{-j\pi\frac{2lN_s}{N}(\epsilon+k_p\delta)}\right\} = \sum_{l=1}^{L}\Re\left\{e^{-j\pi\frac{\alpha_l}{N}}\sum_{k_p\in\mathcal{K}_p}R_{l,k_p}e^{-j\pi\frac{2lN_s}{N}k_p\delta}\right\}. \quad (10)$$

Maximizing (10) with respect to $\boldsymbol{\alpha}$, while δ is fixed, $\delta = \delta_0$, results in the following closed form expression, [25, Equation (17)] $\hat{\alpha}_l(\delta_0) = \frac{N}{\pi}\arg\left\{\sum_{k_p\in\mathcal{K}_p}R_{l,k_p}e^{-j\pi\frac{2lN_s}{N}k_p\delta_0}\right\}$. Plugging $\hat{\alpha}_l(\delta_0)$ back into (10) and maximizing with respect to δ leads to the estimator [25, Equation (18)]:

$$\hat{\delta}_{\mathrm{DML}}^{(\mathrm{uk})} = \underset{\delta\in\mathfrak{R}_\delta}{\operatorname{argmax}}\left\{\sum_{l=1}^{L}\left|\sum_{k_p\in\mathcal{K}_p}R_{l,k_p}e^{-j\pi\frac{2lN_s}{N}k_p\delta}\right|\right\}. \quad (11)$$

Thus, finding $\hat{\delta}_{\mathrm{DML}}^{(\mathrm{uk})}$ requires a search over $\delta\in\mathfrak{R}_\delta$. The work [25] proposed then to use the BLUE to estimate ϵ based on the expanded vector $\hat{\boldsymbol{\alpha}}_{\mathrm{DML}}^{(\mathrm{uk})} \triangleq \left\{\hat{\alpha}_l\left(\hat{\delta}_{\mathrm{DML}}^{(\mathrm{uk})}\right)\right\}_{l=1}^{L}$. Since the BLUE requires the covariance matrix of the estimation errors corresponding to $\hat{\boldsymbol{\alpha}}_{\mathrm{DML}}^{(\mathrm{uk})}$, [25] suggested to use the efficiency property of the MLE, [28, Theorem 7.1], and obtained this matrix by calculating the CRB for the joint estimation of $\boldsymbol{\alpha}$ and δ. We note here that as the estimates $\hat{\alpha}_l\left(\hat{\delta}_{\mathrm{DML}}^{(\mathrm{uk})}\right)$ are clearly suboptimal, it does not imply that $\hat{\boldsymbol{\alpha}}_{\mathrm{DML}}^{(\mathrm{uk})}$ is an efficient estimation of $\boldsymbol{\alpha}$. Thus, using the

respective CRB can only be viewed as an ad-hoc approximation of the required covariance matrix. Furthermore, in [25], the CRB for joint estimation of $\boldsymbol{\alpha}$ and δ was calculated based on the concentrated log-likelihood function (LLF) [25, Equation (16)], instead of using the actual LLF [25, Equation (8)], further reducing the accuracy of the variance expression.

The definition of α_l implies that in the absence of noise, the effective range of this method is approximately $|\epsilon| < \frac{N}{2LN_s}$. In practice, this range can be larger since ϵ can be correctly estimated even if it causes an ambiguity in α_L.

For the case of known CIR, the above decoupling procedure can be repeated. The resulting estimator $\hat{\delta}^{(\mathrm{k})}_{\mathrm{DML}}$ is given by (11) with R_{l,k_p} replaced with $Y_{l,k_p}, \pi\frac{2lN_s}{N}$ replaced with ξ_l, and $\alpha_l = 2lN_s\epsilon$ replaced with $\alpha_l = (N-1+2N_l)\epsilon$. Once the estimate $\hat{\delta}^{(\mathrm{k})}_{\mathrm{DML}}$ is obtained, the BLUE is used for estimating ϵ based on $\hat{\boldsymbol{\alpha}}^{(\mathrm{k})}_{\mathrm{DML}}$. Again, we emphasize that $\hat{\boldsymbol{\alpha}}^{(\mathrm{k})}_{\mathrm{DML}}$, is not necessarily efficient. The effective estimation range in this case is $|\epsilon| < \frac{N}{N-1+2N_L}$.

4. APPROXIMATELY OPTIMAL PER SUBCARRIER ESTIMATION

In this section, we first present a polynomial-based approximate solution for a single dimension maximization problem corresponding to (6). Then, we present a new estimation method that first uses the polynomial approximation to estimate $\rho_{k_p} \triangleq \epsilon + k_p\delta$ for each $k_p \in \mathcal{K}_p$, and then uses the BLUE to estimate ϵ and δ from the per-subcarrier estimates $\hat{\rho}_{k_p}$. We refer to this method as *approximately optimal per subcarrier estimation* (AOPSE).

4.1. Polynomial Approximation for the MLE Cost Function

For $\alpha_l \in \mathfrak{R}, \beta_l \in \mathfrak{C}, l = 1,2,\ldots,L$, consider the maximization problem $\mathrm{argmax}_{x\in\mathfrak{R}}\left\{\sum_{l=1}^{L}\Re\left\{e^{-j\alpha_l x}\cdot\beta_l\right\}\right\}$. Define $\Gamma_0(x) \triangleq \sum_{l=1}^{L}\Re\left\{e^{-j\alpha_l x}\beta_l\right\} = \sum_{l=1}^{L}(\cos(\alpha_l x)\Re\{\beta_l\} + \sin(\alpha_l x)\Im\{\beta_l\})$. The maximizing x must satisfy $\frac{\partial\Gamma_0(x)}{\partial x} = 0$, which can be explicitly written as:

$$\sum_{l=1}^{L}(\alpha_l\cos(\alpha_l x)\Im\{\beta_l\} - \alpha_l\sin(\alpha_l x)\Re\{\beta_l\}) = 0. \quad (12)$$

For $L > 1$, (12) has no closed form solution. However, an approximated solution can be obtained by applying the Taylor expansions (around $x = 0$) of $\sin(\alpha_l x)$ and $\cos(\alpha_l x)$: note that if $|x|$ is small, then the Taylor series is well approximated using a finite number of elements, $M \in \mathfrak{N}$. Plugging these approximations into (12), we obtain that solving (12) amounts to finding the roots of the following polynomial in the variable x:

$$\sum_{n=0}^{M}\frac{(-1)^n}{(2n)!}\sum_{l=1}^{L}\Im\{\beta_l\}\alpha_l^{2n+1}x^{2n} - \sum_{n=0}^{M}\frac{(-1)^n}{(2n+1)!}\sum_{l=1}^{L}\Re\{\beta_l\}\alpha_l^{2n+2}x^{2n+1}. \quad (13)$$

Next, we use this approximation in the AOPSE method.

4.2. AOPSE for Known CIR

Let the range of ρ_{k_p} be denoted by $\mathfrak{R}_\rho(k_p)$. Writing the cost function in the curly brackets in the RHS of (6) in terms of ρ_{k_p} gives $\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\Re\left\{Y_{l,k_p}e^{-j\xi_l\rho_{k_p}}\right\}$. For a specific $k_p \in \mathcal{K}_p$, the MLE of ρ_{k_p} is given by:

$$\hat{\rho}^{(\mathrm{k})}_{k_p,\mathrm{ML}} = \underset{\rho_{k_p}\in\mathfrak{R}_\rho(k_p)}{\mathrm{argmax}}\left\{\sum_{l=1}^{L}\Re\left\{Y_{l,k_p}e^{-j\xi_l\rho_{k_p}}\right\}\right\}. \quad (14)$$

Applying the approximate solution presented in Subsection 4.1, we obtain that $\hat{\rho}^{(\mathrm{k})}_{k_p,\mathrm{ML}}$ is approximately given by one of the roots of the polynomial (in the variable ρ_{k_p}):

$$\sum_{n=0}^{M}\frac{(-1)^n}{(2n)!}\sum_{l=1}^{L}\Im\left\{Y_{l,k_p}\right\}\xi_l^{2n+1}\rho_{k_p}^{2n} - \sum_{n=0}^{M}\frac{(-1)^n}{(2n+1)!}\sum_{l=1}^{L}\Re\left\{Y_{l,k_p}\right\}\xi_l^{2n+2}\rho_{k_p}^{2n+1}, \quad (15)$$

where $M \in \mathfrak{N}$ is a design parameter which controls the tightness of the approximation. Let $\hat{\rho}^{(\mathrm{k})}_{k_p,\mathrm{poly}}$ denote the selected root. Before describing how to solve (15), we consider the possible range of the ρ_{k_p}'s: Equation (14) implies that in absence of noise, an estimation of the effective range of this method is $|\epsilon + k_p\delta| < \frac{N}{N-1+2N_L}$. As stated in Subsection 3.3, in practice, this range can be larger.

The degree of the polynomial in (15) is $2M+1$, thus it has $2M+1$ roots. Some of these roots might be complex, out of which the estimate of ρ_{k_p} has to be selected. We propose to use the Newton-Raphson (NR) method [30], with N_{NR} iterations, for finding $\hat{\rho}^{(\mathrm{k})}_{k_p,\mathrm{poly}}$ explicitly. While the NR method has a low complexity, it might converge to a root different than the root we seek. In order to overcome this drawback, as $\mathfrak{R}_\rho(k_p)$ is set to be symmetric around zero, we apply the iterative method with N_0 different initial conditions in the set $\mathfrak{B}$. For $N_0 = 1$ we set $\mathfrak{B} = \{0\}$. Otherwise we set:

$$\mathfrak{B} = \begin{cases}\left\{\pm\frac{i}{\frac{N_0}{2}+1}\max\left\{\mathfrak{R}_\rho(k_p)\right\}\right\}_{i=1}^{\frac{N_0}{2}} & \text{even } N_0\\ \left\{\pm\frac{i}{\frac{N_0-1}{2}}\max\left\{\mathfrak{R}_\rho(k_p)\right\}\right\}_{i=1}^{\frac{N_0-1}{2}}\cup\{0\} & \text{odd } N_0,\end{cases}$$

and then select the root that maximizes (14) as $\hat{\rho}^{(\mathrm{k})}_{k_p,\mathrm{poly}}$.

Let the vector of all K_p estimates be denoted with $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}} \triangleq \left[\hat{\rho}^{(\mathrm{k})}_{k_1,\mathrm{poly}}, \hat{\rho}^{(\mathrm{k})}_{k_2,\mathrm{poly}}, \ldots, \hat{\rho}^{(\mathrm{k})}_{k_{K_p},\mathrm{poly}}\right]^T$. Define $\mathbb{A}$ to be a $K_p \times 2$ matrix whose elements are $[\mathbb{A}]_{p,1} = 1, [\mathbb{A}]_{p,2} = k_p$, for $p \in \{1, 2, \ldots, K_p\}$, and let $\boldsymbol{\nu}$ be a vector containing the errors in estimating ρ_{k_p}. The relationship between the estimates $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}$ and $\boldsymbol{\theta}$ can thus be written as:

$$\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}} = \mathbb{A}\boldsymbol{\theta} + \boldsymbol{\nu}. \tag{16}$$

The CRB for the estimation of $\boldsymbol{\rho} \triangleq \rho_{k_p}{}_{p=1}^{K_p}$, based on the model (4), is a diagonal matrix, $\mathbb{G}$, whose diagonal elements are:

$$[\mathbb{G}]_{p,p} = \frac{N^2\sigma_\upsilon^2}{2\pi^2|H_{k_p}|^2 \mathrm{F}_{N,L}}, \quad p \in \{1, 2, \ldots, K_p\}. \tag{17}$$

It can be easily verified that for the estimation problem (14) the regularity conditions for efficiency of the MLE, specified in [28, Appendix 7B], hold. Therefore, the MLE (14) is efficient for each k_p. Thus, when the ICI is negligible, $\boldsymbol{\nu}$ can modeled as a zero-mean vector, and its covariance matrix, denoted by $\mathbb{Q}_\nu$, is equal to $\mathbb{G}$. Now, the BLUE for $\boldsymbol{\theta}$ can be computed following [28, Chapter 6]:

$$\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} = \left(\mathbb{A}^H\mathbb{Q}_\nu^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_\nu^{-1}\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}. \tag{18}$$

It thus follows that for negligible ICI, $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ is unbiased and its covariance matrix is given by (Appendix B):

$$\mathbb{Q}^{(\mathrm{k})}_{\mathrm{poly}} = \frac{\sigma_\upsilon^2 N^2}{2\pi^2 \mathrm{F}_{N,L}} \cdot \begin{bmatrix} Q^\star(3) & Q^\star(2) \\ Q^\star(2) & Q^\star(1) \end{bmatrix}. \tag{19}$$

4.3. AOPSE for Unknown CIR

Following similar steps to those applied in Subsection 4.2, the AOPSE method can be applied to the case of unknown CIR. For a *specific* $k_p \in \mathcal{K}_p$, the MLE of ρ_{k_p} is given by:

$$\hat{\rho}^{(\mathrm{uk})}_{k_p,\mathrm{ML}} = \underset{\rho_{k_p}\in\Re_\rho(k_p)}{\mathrm{argmax}} \left\{\sum_{l=1}^{L-1} \Re\left\{R_{l,k_p} e^{-j\pi\frac{2lN_s}{N}\rho_{k_p}}\right\}\right\}. \tag{20}$$

Let $\eta_l = \pi\frac{2lN_s}{N}$. Applying the approximation method presented in Subsection 4.1, we obtain that $\hat{\rho}^{(\mathrm{uk})}_{k_p,\mathrm{ML}}$ is approximately given by one of the roots of the polynomial (in the variable ρ_{k_p}):

$$\begin{aligned}\sum_{n=0}^{M}\frac{(-1)^n}{(2n)!}\sum_{l=1}^{L-1}\Im\left\{R_{l,k_p}\right\}\eta_l^{2n+1}\rho_{k_p}^{2n} - \\ \sum_{n=0}^{M}\frac{(-1)^n}{(2n+1)!}\sum_{l=1}^{L-1}\Re\left\{R_{l,k_p}\right\}\eta_l^{2n+2}\rho_{k_p}^{2n+1}.\end{aligned} \tag{21}$$

Denote the vector of estimated roots by $\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}$. $\boldsymbol{\theta}$ is estimated from $\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}$ via the BLUE. As in Subsection 4.2 ,we have that $\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}} = \mathbb{A}\boldsymbol{\theta} + \boldsymbol{\nu}$, where $\mathbb{A}$ is defined in Subsection 4.2. Here, the CRB for the estimation of $\boldsymbol{\rho}$ is a diagonal matrix, $\mathbb{F}$, whose diagonal entries are:

$$[\mathbb{F}]_{p,p} = \frac{N^2\sigma_\upsilon^2}{2\pi^2|H_{k_p}|^2 \mathrm{J}_{N,L}}, \quad p \in \{1, 2, \ldots, K_p\}. \tag{22}$$

Similarly to Subsection 4.2, it can be verified that the regularity conditions for the estimation problem (20) hold, hence, from the efficiency property of the MLE, when the ICI is negligible, $\boldsymbol{\nu}$ can be modeled as a zero-mean vector, and its covariance matrix, denoted by $\mathbb{Q}_\nu$, is equal to $\mathbb{F}$. Next, we compute the BLUE for $\boldsymbol{\theta}$ and obtain $\hat{\boldsymbol{\theta}}^{(\mathrm{uk})}_{\mathrm{poly}} = \left(\mathbb{A}^H\mathbb{Q}_\nu^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_\nu^{-1}\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}$. As in Subsection 4.2, it follows that $\hat{\boldsymbol{\theta}}^{(\mathrm{uk})}_{\mathrm{poly}}$ is unbiased and its covariance matrix is given by:

$$\mathbb{Q}^{(\mathrm{uk})}_{\mathrm{poly}} = \frac{\sigma_\upsilon^2 N^2}{2\pi^2 \mathrm{J}_{N,L}} \cdot \begin{bmatrix} Q^\star(3) & Q^\star(2) \\ Q^\star(2) & Q^\star(1) \end{bmatrix}. \tag{23}$$

Following the same analysis of the estimation range applied in Subsection 4.2, it follows that the effective range of this method is $|\epsilon + k_p\delta| < \frac{N}{2LN_s}$.

We further note that in order to evaluate $\hat{\boldsymbol{\theta}}^{(\mathrm{uk})}_{\mathrm{poly}}$, the unknown CIR must be estimated. Here, we propose to estimate the CIR based on $\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}$. For unknown CIR, we write $\Gamma(\boldsymbol{\theta})$ in (5) in terms of $\boldsymbol{\rho}$ and as a function of the unknown CIR $\mathbf{H} = \left[H_{k_1}, H_{k_2}, \ldots, H_{k_{K_p}}\right]^T$:

$$\begin{aligned}\Gamma(\boldsymbol{\rho}, \mathbf{H}) = -LK_p \ln\left(\pi\sigma_\nu^2\right) - \\ \frac{1}{\sigma_\nu^2}\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\left|Z_{l,k_p} - e^{j\xi_l\rho_{k_p}}\cdot D_{l,k_p}H_{k_p}\right|^2.\end{aligned}$$

Next, let $X_{l,k_p} = Z_{l,k_p}D^*_{l,k_p}$ and fix $\boldsymbol{\rho} = \hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}$. Then, $\hat{H}_{k_p}\left(\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}\right)$ is obtained by maximizing $\Gamma\left(\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}, \mathbf{H}\right)$:

$$\hat{H}_{k_p}\left(\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}\right) = \frac{1}{L}\sum_{l=1}^{L} X_{l,k_p} e^{-j\xi_l\hat{\rho}^{(\mathrm{uk})}_{k_p,\mathrm{poly}}}, \tag{24}$$

where $\hat{\rho}^{(\mathrm{uk})}_{k_p,\mathrm{poly}}$ is estimated via (21). Finally, we note that using $\hat{H}_{k_p}\left(\hat{\boldsymbol{\rho}}^{(\mathrm{uk})}_{\mathrm{poly}}\right)$ in $\mathbb{F}$, see (22), results in an approximated CRB for the estimation of $\boldsymbol{\rho}$.

4.4. Estimation using partial knowledge of the CIR

In this subsection, we discuss two special cases of partial CIR. First, we consider the case of known CIR gains, $|H_{k_p}|$,

and unknown CIR phases ϕ_{k_p}, and derive the MLE for this scenario. Then, we show that the optimal estimation performance for the case of unknown CIR phases is identical to the optimal estimation performance for the case of unknown CIR. For the case of known CIR phases ϕ_{k_p} and unknown CIR gains $|H_{k_p}|$, we show that the optimal estimation performance is identical to the optimal estimation performance for the case of complete CIR knowledge. This implies that the estimation performance is largely determined by the knowledge of CIR phases, which agrees well with the intuition. Therefore, in order to obtain a good CFO and SFO estimations, the receiver should focus on obtaining accurate estimates of the CIR phases, while the estimates of the CIR gains can be less accurate. We further show that the AOPSE adopts to partial CIR, in contrast to the DML, and in fact it requires only the phases of the CIR coefficients to obtain approximately optimal performance. Lastly, we note that when CIR estimates are only moderately known, then the partial CIR scheme, based on knowledge of only the phases, maybe better than the scheme based on known CIR.

4.4.1. Known gains and unknown phases.

The model (4), stating the channel gains $|H_{k_p}|$ and phases ϕ_{k_p} explicitly, is given by:

$$Z_{l,k_p} = e^{j\xi_l(\epsilon+k_p\delta)} \cdot D_{l,k_p}|H_{k_p}|e^{j\phi_{k_p}} + V_{l,k_p}. \tag{25}$$

In Appendix C.1, we show that the *CRB for estimating $\boldsymbol{\theta}$ based on (25), assuming the CIR phases are unknown while the CIR gains are known, is identical to the CRB for the case of no CIR knowledge at all,* stated in (9). Therefore, the case of completely unknown CIR and the case of unknown CIR phases are equivalent in terms of optimal estimation performance. Next, consider the MLE for $\boldsymbol{\theta}$: We treat the vector of phases $\boldsymbol{\phi} = \left[\phi_{k_1}, \phi_{k_2}, \ldots, \phi_{k_{K_p}}\right]^T$ as a deterministic unknown nuisance parameters vector. The MLE of $\boldsymbol{\phi}$ is found by maximizing the joint LLF of the extended unknown parameter vector $[\boldsymbol{\theta}^T, \boldsymbol{\phi}^T]^T$. In this case, from the invariance property of the MLE [28, Thm. 7.2], the MLE of $\boldsymbol{\phi}$, for a fixed $\boldsymbol{\theta}$, is found to be:

$$\hat{\phi}^{(\text{partial})}_{k_p,\text{ML}}(\boldsymbol{\theta}) = \arg\left\{\sum_{l=1}^{L} X_{l,k_p} e^{-j\xi_l(\epsilon+k_p\delta)}\right\}. \tag{26}$$

Plugging (26) into the LLF (5), we obtain the MLE for $\boldsymbol{\theta}$ (Appendix C.2):

$$\hat{\boldsymbol{\theta}}^{(\text{phase})}_{\text{ML}} = \underset{\boldsymbol{\theta}\in\Re_\epsilon\times\Re_\delta}{\text{argmax}} \left\{\sum_{k_p\in\mathcal{K}_p} |H_{k_p}| \times \left|\sum_{l=1}^{L} X_{l,k_p} e^{-j\xi_l(\epsilon+k_p\delta)}\right|\right\}. \tag{27}$$

In contrast to the MLE for unknown CIR, here the decoupling procedure of Subsection 3.3 cannot be applied because the MLE in (27) is stated in terms of the L_1-norm. Therefore, the MLE for the case of unknown phase cannot be simplified using the same approach taken for unknown CIR. However, the AOPSE approach presented in Subsection 4.1 can be adapted to the case of unknown CIR phases. Hence, for a specific subcarrier $k_p \in \mathcal{K}_p$ we have

$$\begin{aligned}\hat{\rho}^{(\text{phase})}_{k_p,\text{ML}} &= \underset{\rho_{k_p}\in\Re_\rho(k_p)}{\text{argmax}} \left\{|H_{k_p}| \left|\sum_{l=1}^{L} X_{l,k_p} e^{-j\xi_l N\rho_{k_p}}\right|\right\} \\ &= \underset{\rho_{k_p}\in\Re_\rho(k_p)}{\text{argmax}} \left\{\sum_{l=1}^{L-1} \Re\left\{R_{l,k_p} e^{-j\pi\frac{2lN_g}{N}\rho_{k_p}}\right\}\right\},\end{aligned} \tag{28}$$

where $|H_{k_p}|$ is omitted because it is constant for a fixed k_p. Note that (28) and (20) are identical, therefore $\boldsymbol{\theta}$ can be estimated via the AOPSE derived in Subsection 4.3.

4.4.2. Unknown gains and known phases.

Following steps similar to those detailed in Appendix C.1, it can be shown that the CRB for estimating $\boldsymbol{\theta}$ based on (25), assuming the CIR gains are unknown while the CIR phases are known, is identical to the CRB for the case of completely known CIR, see (7). Therefore, the case of completely known CIR and the case of unknown CIR gains are equivalent in terms of optimal estimation performance. Next, consider the MLE for $\boldsymbol{\theta}$, and let $\lambda_{k_p} = |H_{k_p}|$. We treat the vector of gains $\boldsymbol{\lambda} = \left[\lambda_{k_1}, \lambda_{k_2}, \ldots, \lambda_{k_{K_p}}\right]^T$ as a deterministic unknown nuisance parameters vector. The MLE of $\boldsymbol{\lambda}$ is the vector $\tilde{\boldsymbol{\lambda}}$ that maximizes the joint LLF of the extended unknown parameter vector $[\boldsymbol{\theta}^T, \boldsymbol{\lambda}^T]^T$ (Appendix C.3):

$$\hat{\lambda}^{(\text{partial})}_{k_p,\text{ML}}(\boldsymbol{\theta}) = \frac{1}{L}\sum_{l=1}^{L} \Re\left\{X_{l,k_p} e^{-j\phi_{k_p}} e^{-j\xi_l(\epsilon+k_p\delta)}\right\}. \tag{29}$$

Plugging (29) into the LLF (5), we obtain:

$$\hat{\boldsymbol{\theta}}^{(\text{gain})}_{\text{ML}} = \underset{\boldsymbol{\theta}\in\Re_\epsilon\times\Re_\delta}{\text{argmax}} \left\{\sum_{k_p\in\mathcal{K}_p} \left(\sum_{l=1}^{L} \Re\left\{X_{l,k_p} e^{-j\phi_{k_p}} e^{-j\xi_l(\epsilon+k_p\delta)}\right\}\right)^2\right\}. \tag{30}$$

Similarly to Subsection 4.4.1, the decoupling procedure cannot be applied, but the AOPSE approach presented in Subsection 4.1 can be adapted to the case of unknown CIR gains.

This indicates another advantage of the AOPSE: it applies to the case of partial CIR, this is in contrast to the DML presented in Subsection 3.3.

5. SIMULATIONS AND NUMERICAL PERFORMANCE EVALUATIONS

5.1. Simulation parameters and review of previous studies

The performance of the estimation schemes presented in this work is evaluated for a system with parameters compatible with the IEEE 802.16 standard for fixed OFDM-based wireless metro access network (MAN), operating at the 5 GHz band [2, Ch. 8.3] (due to the similarity in configuration, the simulation scenario is also relevant to LTE [3]), see also [25]: we assume DFT of length $N = 256$ and CP of length $N_g = 16$. The sampling period is $T = 87.5\,ns$, and the subcarrier spacing is 44.6 KHz. Furthermore, $N_D = 200$ modulated symbols (information and pilots) are used per OFDM symbol; these symbols are located at the subcarriers with frequency indices $\{\pm 100, \pm 99, \ldots, \pm 1\}$, where eight pilots are located at the subcarriers whose frequency indices are $\{\pm 88, \pm 63, \pm 38, \pm 13\} \triangleq \mathcal{K}_p$, while the remaining subcarriers are dedicated for data. We consider a frequency selective slow Rayleigh fading channel with an impulse response of length $L_{\text{ch}} = 12$, where each channel tap is modeled as a zero-mean statistically independent complex Normal RV, with an exponential power delay profile given by $\mathrm{E}\left\{|h_i|^2\right\} = e^{-\frac{2i}{L_{\text{ch}}}} \Big/ \sum_{i=0}^{L_{\text{ch}}} e^{-\frac{2i}{L_{\text{ch}}}}$, $i = 0, 1, \ldots, L_{\text{ch}}-1$. As stated in Section 2, the channel remains constant over an OFDM frame. In the simulations, a different channel realization is generated for each OFDM frame, except for simulations that include a comparison with the CRB, where the same channel realization[‡] is used for all OFDM frames. Note that the CRBs given in (7) and (9) hold for a given CIR. When the CIR is a random vector, a hybrid CRB [31] should be used, however, this is out of the scope of this work.

The CFO is taken as 45.7 ppm, which for the above parameters corresponds to $\epsilon = 0.02$ and the SFO is taken as 100 ppm, which corresponds to $\delta = 10^{-4}$. CFO and SFO estimation is done using a frame consisting of $L = 8$ OFDM symbols. For the DML the coarse search is executed over $N_{\text{srch}} = 16$ equidistant values of δ in the range $|\delta| < 5 \cdot 10^{-4}$. Furthermore, we set $N_{\text{expan}} = 20$ to be the interpolation expansion factor. The residual CFO ϵ is in the range $|\epsilon| < 0.05$. For the AOPSE of Section 4 we use Taylor series of order $M = 3$. The maximal number of iterations for the NR method is set to $N_{\text{NR}} = 5$, and the number of initial conditions for the NR method is set to $N_0 = 2$. The symbols at the pilot subcarriers are uniformly selected from a BPSK constellation, while the data subcarriers are modulated with equiprobable QPSK symbols.

Throughout this section, we compare our results to the estimation methods presented in [20,23,25] and [17]. The methods of [20] and [23] do not rely on known pilots or CIR. Instead, it is assumed that the same pilots are transmitted in two consecutive OFDM symbols. Then, the phase differences are calculated via $\check{Z}_{l,k_p} = Z_{l+1,k_p} Z^*_{l,k_p}$. The work [20] estimates ϵ and δ via $\sum_{k_p \in \mathcal{K}_p} \check{Z}_{l,k_p}$; this method is denoted by SFFM (Speth-Fechtel-Fock-Meyr). The work [23] uses the LS estimator [28, Chapter 8] to estimate ϵ and δ out of $\left\{\check{Z}_{l,k_p}\right\}_{k_p \in \mathcal{K}_p}$; this method is denoted by LC (Liu-Chong). Note that these methods use only two consecutive OFDM symbols. For fairness of comparison, we extend the schemes of [20] and [23] to $L > 2$ OFDM symbols via averaging the different estimations. The work in [17] assumes that the CIR is known and approximates the term $e^{j\xi_l(\epsilon + k_p\delta)}$ in (4) by $1 + j\xi_l(\epsilon + k_p\delta)$, thus, the LLF in (5) becomes linearly dependent on ϵ and δ. Then, by setting the derivative of this linear model to zero, [17] obtains a closed form approximate MLEs for ϵ and δ. We denote this method by HBC (Häring-Bieder-Czylwik). Finally, the reduced complexity estimator (RCE), presented in [25], is a two-stage linear method in which the receiver first estimates the quantities $\left\{\rho_{k_p}\right\}_{k_p \in \mathcal{K}_p}$ using the linear estimation method established in [32]. Then, based on $\left\{\rho_{k_p}\right\}_{k_p \in \mathcal{K}_p}$, the BLUE is used for estimating ϵ and δ.

5.2. Simulation results

First, we present the mean-squared-error (MSE) of the different methods. The MSE is defined as $\text{MSE}(\delta) \triangleq \mathrm{E}\{(\delta - \hat{\delta})^2\}$ and $\text{MSE}(\epsilon) \triangleq \mathrm{E}\{(\epsilon - \hat{\epsilon})^2\}$. Figure 1 depicts $\text{MSE}(\delta)$ and $\text{MSE}(\epsilon)$ versus SNR for the case of known CIR. In this case, the relevant methods are the DML, AOPSE, and HBC. The CRB for this scenario is depicted as a reference. It can be observed that for the case of known CIR, as expected, both the DML and AOPSE approach the CRB for low and medium SNR conditions. We also observe that this does not hold at high SNR because the ICI is no longer negligible. Lastly, we note that the HBC performs very poorly. This follows from the fact that for $L = 8$, the approximation $e^{j\xi_l(\epsilon + k_p\delta)} \approx 1 + j\xi_l(\epsilon + k_p\delta)$, applied in [17], does not hold.

Next, Figure 2 depicts $\text{MSE}(\delta)$ and $\text{MSE}(\epsilon)$ vs. SNR for the case of unknown CIR. In this case, the relevant methods are LC, RCE, SFFM, DML and AOPSE. It can be observed that for medium and high SNR conditions, RCE, AOPSE and DML achieve approximately the same performance. Similarly to the case of known CIR, when the SNR is high, the CRB is not attainable by these schemes. However, when the SNR is low, the RCE suffers from severe degradation in the estimation accuracy. AOPSE and DML, on the other hand, approach the CRB for both low as well as for medium SNR conditions, at the cost of higher computational complexity. Figure 2 also shows

[‡]The channel taps are given by $\mathbf{H} = [-0.4833 + 0.0683j, -0.2686 + 0.4397j, 0.1396 - 0.3578j, -0.1589 - 0.1945j, -0.0659 - 0.0590j, -0.2321 + 0.0699j, 0.0698 - 0.0741j, -0.1506 - 0.1815j, -0.2894 - 0.0719j, -0.2140 + 0.2358j, 0.0434 - 0.0203j, -0.0280 - 0.0261j]^T$. Note that $|\mathbf{H}|^2 = 0.99994$.

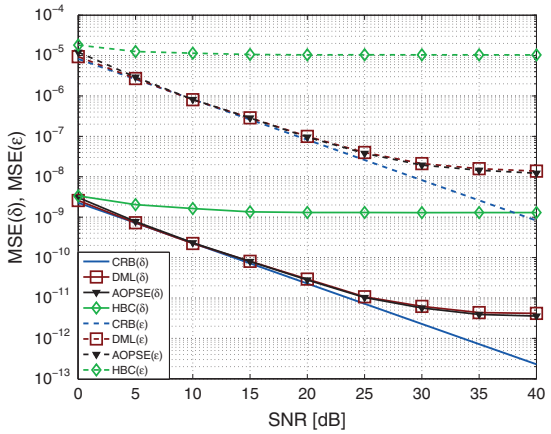

Figure 1. MSE(δ) and MSE(ϵ) versus SNR for known CIR with $\epsilon = 0.02, \delta = 10^{-4}$, and $L = 8$.

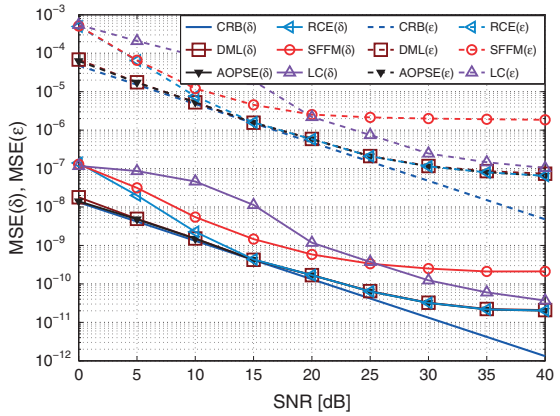

Figure 2. MSE(δ) and MSE(ϵ) versus SNR for unknown CIR with $\epsilon = 0.02, \delta = 10^{-4}$, and $L = 8$.

that the performance of LC and SFFM are worse than the performance of the other methods.

Next, we examine the effect of knowledge of the CIR on the estimation performance by comparing the estimation accuracy and range in known CIR with unknown CIR. This is illustrated in Figures 3 and 4. Figure 3 depicts the average estimated ϵ versus the true value of ϵ, for known CIR with $\delta = 10^{-4}$ and SNR $= 15$ dB. The `Unbiased` line is presented as a reference. Figure 3 agrees with our analysis that shows that the DML, for the case of known CIR, is unbiased for $|\epsilon| < 0.055$, while for the case of unknown CIR, the DML is unbiased for $|\epsilon| < 0.058$. Furthermore, it can be observed that the practical estimation range of the AOPSE is larger than the one for the DML, as indicated in Section 4. Additionally, note that in Figure 3 , the estimation range in the case of unknown CIR is larger than in the case of known CIR. This holds in particular for AOPSE. The MSE comparison for known versus unknown CIR is depicted in Figure 4. Observe that the gain in MSE due to knowledge of the CIR can be about 7 dB. This indicates a tradeoff for the methods derived here: while CIR knowledge significantly improves the estimation accuracy over unknown CIR, it limits the estimation range.

Finally, the effect of the frame length L on the estimation performance is demonstrated in Figure 5, which depicts MSE(ϵ) as a function of L for unknown CIR at SNR $= 7$ dB, $\epsilon = 0.02$, and $\delta = 10^{-4}$. From Figure 5, it can be observed that for unknown CIR at SNR $= 7$ dB, DML and AOPSE significantly outperform RCE, SFFM and LC, for the entire range of L tested.

5.3. Additional important system aspects

5.3.1. Continuous wave interference.

We note that as in this work we considered estimation schemes based on pilot tones, then the proposed algorithms as well as all previously proposed schemes are not affected by continuous wave interference (CWI), which is not present in the pilot tones. Yet, these schemes are susceptible to CWI that affects one or more pilot subcarriers. Such CWI may significantly degrade performance of all pilot-based schemes due to inducing a large bias and

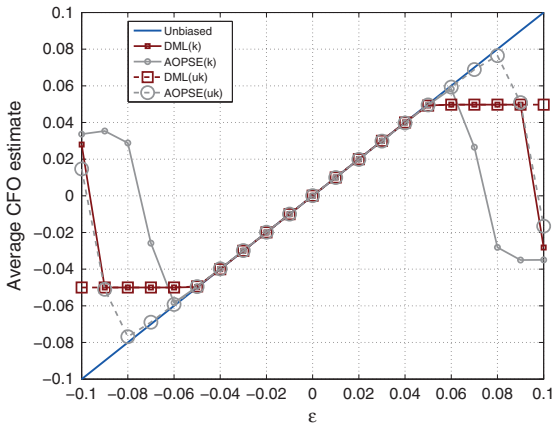

Figure 3. E$\{\hat{\epsilon}\}$ versus ϵ for $\delta = 10^{-4}$, $L = 8$, and SNR $= 15$ dB.

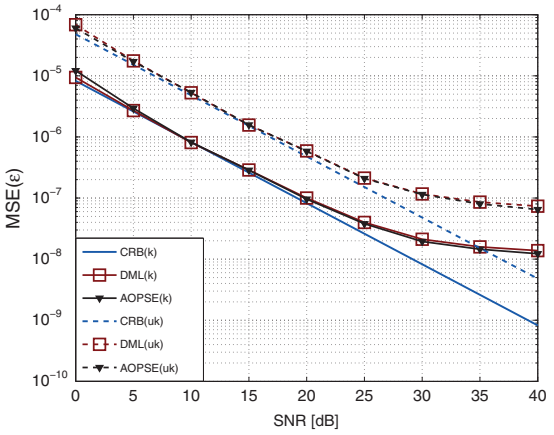

Figure 4. MSE(ϵ) versus SNR for $\epsilon = 0.02, \delta = 10^{-4}$, and $L = 8$.

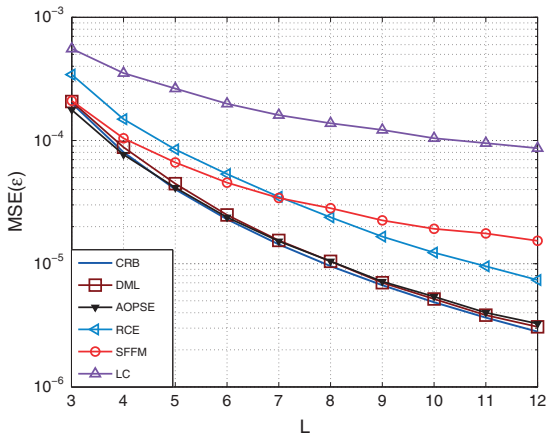

Figure 5. MSE(ϵ) versus L for unknown CIR with $\epsilon = 0.02$, $\delta = 10^{-4}$, and SNR = 7 dB.

increasing the variance of the estimators. We note however, that the impact of a CWI located in one of the pilot subcarriers can be partially mitigated by a simple modification to the estimation scheme, for example, by applying energy detector to the pilot tones and excluding from the estimator all pilot tones suspected of suffering from CWI. Then, the AOPSE is applied using only the pilots subcarriers in which no CWI was detected.

5.3.2. Bit error rate.

For the simulated system, CFO and SFO of $\epsilon = 0.02$ and $\delta = 10^{-4}$, respectively, correspond to clock offsets of about 50 ppm for CFO and about 100 ppm for the SFO. These values are reasonable for the outcome of the initial coarse synchronization [25]. Yet, from additional simulations we carried out, we conclude that these values do not allow functional system operation, as if left uncorrected the resulting uncoded bit error rate (BER) is very high, for example, for the case of known CIR, an uncoded BER of 0.5 is measured at all SNR values in the range of 0 dB to 40 dB. When applying CFO and SFO estimation and correction with the DML or with the AOPSE schemes, then the uncoded BER approaches the case of no CFO and SFO at all, and decreases linearly (on the log-log scale) with the SNR. We conclude that good SFO and CFO estimation schemes are essential for proper operation of the system.

5.3.3. Different cyclic prefix lengths.

From (4), it follows that the length of the CP affects the received signal via ξ_l which, for given ϵ and δ, determines the phase drift in each of the pilot subcarriers. Typical lengths of CP, e.g., in IEEE 802.16, are $\frac{N}{16}$ (used in the present simulations), $\frac{N}{8}$ and $\frac{N}{4}$. After repeating our simulations for $N_g = \frac{N}{8} = 32$ and for $N_g = \frac{N}{4} = 64$, we conclude that the AOPSE scheme approaches the CRB in low-to-medium SNRs, also when $N_g = 32$ and $N_g = 64$.

5.4. Computational complexity

5.4.1. Computational complexity of the DML.

A detailed computational complexity analysis for the DML is provided in Appendix D.1. Here, we provide only an outline of this analysis for the case of known CIR. The DML in Subsection 3.3 consists of two stages: evaluating (11) (for the case of known CIR) via a search and estimating ϵ from the vector $\boldsymbol{\alpha}$ via the BLUE. The search is executed in two steps: a coarse exhaustive search over N_{srch} equidistant values of δ in the range $\delta \in \Re_\delta$ is followed by a cubic spline interpolation with expansion factor N_{expan}. The computational complexity of the coarse exhaustive search is $N_{\text{srch}}(3LK_p + L - 1)$ while the computational complexity of the cubic spline interpolation is $7N_{\text{srch}}N_{\text{expan}} + 2N_{\text{srch}} + 114$, see [33]. Next, estimating ϵ from the vector $\hat{\boldsymbol{\alpha}}$ via the BLUE has a computational complexity equal to $8L^2 + 3LK_p + 6L + 5K_p - 3$. Therefore, the computational complexity of the DML in Subsection 3.3 is equal to: $\mathscr{C}_{\text{DML}} = 3N_{\text{srch}}LK_p + 7N_{\text{srch}}N_{\text{expan}} + N_{\text{srch}}L + 3L^2 + 3LK_p + N_{\text{srch}} + 7L + 5K_p + 115$.

5.4.2. Computational complexity of the AOPSE in subsection 4.2.

A detailed computational complexity analysis for the AOPSE is provided in Appendix D.2. Here, we provide only an outline of this analysis. The AOPSE method first applies the polynomial approximation for each subcarrier, and then uses the BLUE via (18). The computational complexity of evaluating the coefficients of the polynomial (15) is $4LK_p(M+1) - K_p$, while the complexity of finding the desired root via the NR method (for N_0 initial conditions) is $K_p(2M + N_0(10N_{\text{NR}}M - 3N_{\text{NR}} - 2M + 3L + 3))$. Finally, estimating $\boldsymbol{\theta}$ via the BLUE (18) has a computational complexity of $16K_p + 2$. Therefore, the computational complexity of the AOPSE is equal to: $\mathscr{C}_{\text{AOPSE}} = K_p(4L(M+1) + 2M + 15 + N_0(10N_{\text{NR}}M - 3N_{\text{NR}} - 2M + 3L + 3)) + 2$.

5.4.3. Comparison of computational complexity.

As the DML and the AOPSE have approximately the same performance, to give a complete comparison we evaluated the computational complexity of both methods, for the case of known CIR and for the simulation parameters detailed in Subsection 5.1. We obtained that the computational complexity of the DML, presented in Subsection 3.3, is $\mathscr{C}_{\text{DML}} = 6051$ while the computational complexity of the AOPSE, presented in Subsection 4.2, is $\mathscr{C}_{\text{AOPSE}} = 3690$. It can be observed that while the AOPSE achieves the same estimation accuracy as the DML, it requires only about 60% of the computational complexity. We further note that the AOPSE maintains a tradeoff between N_0 and N_{NR}. For example, increasing N_0 from 2 to 3 and decreasing N_{NR} from 5 to 4, the performance become negligibly worst, while complexity increases to 4290 flops.

6. CONCLUSIONS

In this work, we studied the problem of estimating the residual CFO and SFO based on known pilots, in an OFDM-based system for which the CIR is approximately constant over an OFDM frame. Three cases of CIR knowledge at the receiver were considered: full knowledge, no-knowledge, and partial knowledge. We first proposed an approximate MLE through the Taylor expansion. This replaces the MLE search with the problem of finding the roots of a polynomial, which can be solved effectively using numerical methods. This method was used to obtain a polynomial in terms of a linear combination of the CFO and SFO, which was then combined with the BLUE to form an estimate, referred to as AOPSE, which does not require an exhaustive search. The AOPSE was shown to achieve similar performance as the DML but at a significantly lower computational complexity.

For the case of partial CIR knowledge, we showed that the case of known CIR gains and unknown CIR phases is equivalent, in terms of the optimal performance (i.e., the CRB) to completely unknown CIR, while the case of known phases and unknown gains is equivalent to completely known CIR. This implies that in order to obtain good estimations of the CFO and SFO, one should focus on estimating the CIR phases, while the impact of the CIR gains on the estimation is significantly lower. Furthermore, we explained why the AOPSE is robust to the knowledge of the channel gains while the DML is not.

APPENDIX A: DERIVATION OF THE CRB FOR KNOWN CIR

The received signal (4) can be written in a matrix form: $\mathbf{Z} = \mathbb{B}(\epsilon,\delta)\mathbb{D}\mathbf{H} + \mathbf{V}$, where $\mathbf{Z} \triangleq \Big[Z_{1,k_1}, Z_{1,k_2}, \ldots, Z_{1,k_{K_p}}, \; Z_{2,k_1}, \ldots, Z_{L,k_{K_p}} \Big]^T$ is an $LK_p \times 1$ vector and $\mathbf{V}$ is defined in the same manner using $\{V_{l,k_p}\}_{l=1,p=1}^{L,K_p}$; $\mathbb{B}(\epsilon,\delta)$ is an $LK_p \times LK_p$ diagonal matrix with the entries $\left\{e^{j\xi_l(\epsilon+k_1\delta)}, e^{j\xi_l(\epsilon+k_2\delta)}, \ldots, e^{j\xi_l(\epsilon+k_{K_p}\delta)}\right\}$; $\mathbb{D}$ is an $LK_p \times LK_p$ diagonal matrix with the entries $\left\{D_{1,k_1}, D_{1,k_2}, \ldots, D_{1,k_{K_p}}, D_{2,k_1}, \ldots, D_{L,k_{K_p}}\right\}$; and $\mathbf{H} = [\underbrace{\check{\mathbf{H}}, \check{\mathbf{H}}, \ldots, \check{\mathbf{H}}}_{L \text{ times}}]^T$, $\check{\mathbf{H}} = \left[H_{k_1}, H_{k_2}, \ldots, H_{k_{K_p}}\right]$ is an $LK_p \times 1$ vector. Treating $\mathbb{B}(\epsilon,\delta)$, $\mathbf{H}$ and $\mathbb{D}$ as fixed unknowns, $\mathbf{Z}$ is the realization of a circularly symmetric complex Normal vector with unknown mean $\boldsymbol{\mu} = \mathbb{B}(\epsilon,\delta)\mathbb{D}\mathbf{H}$ and known covariance matrix $\sigma_v^2\mathbb{I}_{LK_p}$. Since the matrix $\mathbb{B}(\epsilon,\delta)$ is differentiable in ϵ and δ, the Fisher information matrix (FIM), associated with the unknown vector $\boldsymbol{\theta}$ is given by [28, Chapter 3]:

$$\mathbb{J} = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}}\right\} = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

where, a, b and c are given by:

$$a = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\epsilon}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\epsilon}\right\} = \frac{2\pi^2\mathrm{F}_{N,L}}{\sigma_v^2N^2}\sum_{k_p\in\mathcal{K}_p}|H_{k_p}|^2,$$

$$b = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\epsilon}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\delta}\right\} = \frac{2\pi^2\mathrm{F}_{N,L}}{\sigma_v^2N^2}\sum_{k_p\in\mathcal{K}_p}k_p|H_{k_p}|^2,$$

$$c = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\delta}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\delta}\right\} = \frac{2\pi^2\mathrm{F}_{N,L}}{\sigma_v^2N^2}\sum_{k_p\in\mathcal{K}_p}k_p^2|H_{k_p}|^2.$$

Recalling that $Q(q) = \sum_{k_p\in\mathcal{K}_p}k_p^{q-1}|H_{k_p}|^2, q = 1,2,3$, and that $Q^\star(j) = \frac{Q(j)}{Q(1)Q(3)-Q^2(2)}$, we arrive at:

$$\mathbb{J}^{-1} = \frac{\sigma_v^2{}^2}{N}\,2\pi^2\mathrm{F}_{N,L}\cdot\begin{bmatrix} Q^\star(3) & Q^\star(2) \\ Q^\star(2) & Q^\star(1) \end{bmatrix}. \tag{A.1}$$

The CRBs for the estimation of ϵ and δ are the diagonal elements of $\mathbb{J}^{-1}$, stated in (7).

APPENDIX B: MEAN AND COVARIANCE MATRIX OF $\hat{\boldsymbol{\theta}}^{(\mathrm{K})}_{\mathrm{POLY}}$ IN SUBSECTION 4.2

Plugging (16) into (18) we have:

$$\begin{aligned}\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} &= \left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_v^{-1}\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}} \\ &= \boldsymbol{\theta} + \left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_v^{-1}\boldsymbol{v}.\end{aligned} \tag{B.1}$$

Taking the expectation, and using $\mathrm{E}\{\boldsymbol{v}\} = 0$, we have $\mathrm{E}\left\{\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}\right\} = \boldsymbol{\theta}$, and therefore $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ is unbiased. Next, $\mathbb{Q}^{(\mathrm{k})}_{\mathrm{poly}}$, the covariance matrix of $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$, can be obtained as:

$$\begin{aligned}\mathbb{Q}^{(\mathrm{k})}_{\mathrm{poly}} &= \mathrm{E}\left\{\left(\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} - \boldsymbol{\theta}\right)^H\right\} \\ &= \left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_v^{-1}\mathrm{E}\left\{\boldsymbol{v}\boldsymbol{v}^H\right\}\times \\ &\qquad \left(\mathbb{Q}_v^{-1}\right)^H\mathbb{A}\left(\left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1}\right)^H \\ &= \left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1}.\end{aligned} \tag{B.2}$$

Now, plugging $\mathbb{A}$, defined in Section 4, $\mathbb{Q}_v$ whose elements are given in (17), and explicitly calculating the elements of $\mathbb{A}^H\mathbb{Q}_v^{-H}\mathbb{A}$, we obtain that $\left(\mathbb{A}^H\mathbb{Q}_v^{-1}\mathbb{A}\right)^{-1} = \mathbb{J}^{-1}$ given in (A.1).

APPENDIX C: DETAILED DERIVATIONS FOR PARTIAL CIR IN SUBSECTION 4.4

C.1. CRB for Known Gains and Unknown Phases

The model for the received signal (25) can be written in a matrix form: $\mathbf{Z} = \mathbb{B}(\epsilon,\delta)\mathbb{D}\mathbb{H}\boldsymbol{\phi} + \mathbf{V}$, where $\mathbf{Z}, \mathbf{V}, \mathbb{B}(\epsilon,\delta)$ and $\mathbb{D}$ are defined as in Appendix A, $\mathbb{H}$ is an $LK_p \times LK_p$ diagonal matrix with the entries $[\underbrace{\mathbf{H},\mathbf{H},\ldots,\mathbf{H}}_{L\,\text{times}}], \mathbf{H} = [|H_{k_1}|,|H_{k_2}|,\ldots,|H_{k_{K_p}}|]$, and $\boldsymbol{\phi} = [\underbrace{\check{\boldsymbol{\phi}},\check{\boldsymbol{\phi}},\ldots,\check{\boldsymbol{\phi}}}_{L\,\text{times}}]^T, \check{\boldsymbol{\phi}} = \left[e^{j\phi_{k_1}}, e^{j\phi_{k_2}},\ldots,e^{j\phi_{k_{K_p}}}\right]$ is an $LK_p \times 1$ vector representing the unknown phases. Treating $\mathbb{B}(\epsilon,\delta), \mathbf{H}$ and $\mathbb{D}$ as fixed unknowns, $\mathbf{Z}$ is the realization of a circularly symmetric complex Normal vector with unknown mean $\boldsymbol{\mu} = \mathbb{B}(\epsilon,\delta)\mathbb{D}\mathbb{H}\boldsymbol{\phi}$ and known covariance matrix $\sigma_{\upsilon}^2\mathbb{I}_{LK_p}$. Since the matrix $\mathbb{B}(\epsilon,\delta)$ is differentiable in ϵ and δ, the FIM associated with the unknown parameters vector $\boldsymbol{\theta} = [\boldsymbol{\phi}^T,\epsilon,\delta]^T$ is given by:

$$\mathbb{J}^{(\text{phase})} = \frac{2}{\sigma_{\upsilon}^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\boldsymbol{\theta}}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\boldsymbol{\theta}}\right\} = \begin{bmatrix}\mathbb{A} & \mathbb{M}\\ \mathbb{M}^T & \mathbb{P}\end{bmatrix},$$

where $\mathbb{A}$ is a $K_p \times K_p$ diagonal matrix with the entries:

$$[\mathbb{A}]_{p,p} = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\phi_{k_p}}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\phi_{k_p}}\right\} = \frac{2L|H_{k_p}|^2}{\sigma_v^2}.$$

$\mathbb{M}$ is a $K_p \times 2$ matrix, where the entries of its first column are given by:

$$[\mathbb{M}]_{p,1} = \frac{2}{\sigma_{\upsilon}^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\phi_{k_p}}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\epsilon}\right\} = \frac{2\pi|H_{k_p}|^2}{\sigma_{\upsilon}^2 N}\sum_{l=1}^{L}(N-1+2N_l),$$

and the entries of its second column are given by:

$$[\mathbb{M}]_{p,2} = \frac{2}{\sigma_v^2}\Re\left\{\frac{\partial\boldsymbol{\mu}^H}{\partial\phi_{k_p}}\cdot\frac{\partial\boldsymbol{\mu}}{\partial\delta}\right\} = \frac{2\pi|H_{k_p}|^2 k_p}{\sigma_v^2 N}\sum_{l=1}^{L}(N-1+2N_l).$$

$\mathbb{P}$ is equal to the matrix $\mathbb{J}$ given in (A.1). The Schur complement (see [34, Appendix C.4]) of $\mathbb{P}$ with respect to $\mathbb{J}^{(\text{phase})}$ is given by $\mathbb{S} = (\mathbb{P} - \mathbb{M}^T\mathbb{A}^{-1}\mathbb{M})^{-1}$. The CRBs for the estimation of ϵ and δ are obtained as the diagonal entries of $\mathbb{S}$. From the explicit expressions for $\mathbb{A}$ and $\mathbb{M}$, we obtain that $\mathbb{M}^T\mathbb{A}^{-1}\mathbb{M}$ is equal to:

$$\frac{2\pi^2\left(\sum_{l=1}^{L}(N-1+2N_l)\right)^2}{\sigma_{\upsilon}^2 LN^2}\begin{bmatrix}Q(1) & Q(2)\\ Q(2) & Q(3)\end{bmatrix},$$

and we conclude that the diagonal entries of $\mathbb{S}$ are stated in (9).

C.2. Derivation of (27)

Plugging (26) into (6), and defining $X_{l,k_p} = Z_{l,k_p}D_{l,k_p}^*$, we have $\hat{\boldsymbol{\theta}}_{\text{ML}}^{(\text{phase})}$ in (C.1), at the bottom of the page.

C.3. Derivation of (29)

Using $H_{k_p} = \lambda_{k_p}e^{j\phi_{k_p}}$ in the LLF (5), we arrive at the MLE for $\boldsymbol{\lambda}$ given $\boldsymbol{\theta} = \boldsymbol{\theta}_0, \hat{\boldsymbol{\lambda}}_{\text{ML}}^{(\text{partial})}(\boldsymbol{\theta}_0)$, given in (C.2) at the bottom of the page. By differentiating the expression in the curly brackets in (C.2), with respect to λ_{k_p}, we obtain $\sum_{l=1}^{L}\left\{2\lambda_{k_p} - 2\Re\left\{Z_{l,k_p}D_{l,k_p}^*e^{-j\phi_{k_p}}e^{-j\xi_l(\epsilon_0+k_p\delta_0)}\right\}\right\}$. As the second derivative is positive, the MLE for λ_{k_p} can be written as:

$$\hat{\lambda}_{k_p,\text{ML}}^{(\text{partial})}(\tilde{\boldsymbol{\theta}}) = \frac{1}{L}\sum_{l=1}^{L}\Re\left\{X_{l,k_p}e^{-j\phi_{k_p}}e^{-j\xi_l(\epsilon_0+k_p\delta_0)}\right\}.$$

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\text{ML}}^{(\text{phase})} &= \underset{\boldsymbol{\theta}\in\Re_\epsilon\times\Re_\delta}{\text{argmax}}\left\{\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\Re\left\{Z_{l,k_p}e^{-j\xi_l(\epsilon+k_p\delta)}D_{l,k_p}^*|H_{k_p}|e^{-j\hat{\phi}_{k_p,\text{ML}}^{(\text{partial})}(\boldsymbol{\theta})}\right\}\right\}\\ &= \underset{\boldsymbol{\theta}\in\Re_\epsilon\times\Re_\delta}{\text{argmax}}\left\{\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\Re\left\{X_{l,k_p}|H_{k_p}|e^{-j\xi_l(\epsilon+k_p\delta)}\cdot e^{-j\arg\left\{\sum_{m=1}^{L}X_{m,k_p}e^{-j\xi_m(\epsilon+k_p\delta)}\right\}}\right\}\right\}\\ &= \underset{\boldsymbol{\theta}\in\Re_\epsilon\times\Re_\delta}{\text{argmax}}\left\{\sum_{k_p\in\mathcal{K}_p}|H_{k_p}|\left|\sum_{l=1}^{L}X_{l,k_p}e^{-j\xi_l(\epsilon+k_p\delta)}\right|\right\}\end{aligned} \tag{C.1}$$

$$\hat{\boldsymbol{\lambda}}_{\text{ML}}^{(\text{partial})}(\boldsymbol{\theta}_0) = \underset{\boldsymbol{\lambda}\in(\Re^+)^{K_p}}{\text{argmin}}\left\{\sum_{l=1}^{L}\sum_{k_p\in\mathcal{K}_p}\left|Z_{l,k_p} - e^{j\xi_l(\epsilon_0+k_p\delta_0)}\cdot D_{l,k_p}\lambda_{k_p}e^{j\phi_{k_p}}\right|^2\right\} \tag{C.2}$$

APPENDIX D: DETAILED COMPLEXITY ANALYSIS

The computational complexity is defined as the number of flops needed to complete each calculation. Here, we assume that the CIR is known. Similar results can be obtained for the case of unknown CIR. We also assume that the sequence $\{Z_{l,k_p}\}_{l=1,p=1}^{L,K_p}$, is available at the receiver, and that the symbols at the pilots subcarriers, that is, k_p, are selected from a BPSK constellation [2].

D.1. Computational complexity analysis for the DML of Subsection 3.3

Recall that the DML for the case of known CIR consists of two steps: (1) estimating $\hat{\delta}_{\text{DML}}^{(\text{k})}$ via a search which is executed in two steps: coarse exhaustive search followed by a fine search via a cubic spline interpolation and (2) estimating $\hat{\epsilon}_{\text{DML}}^{(\text{k})}$ via the BLUE based on the vector $\boldsymbol{\alpha}\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)$.

***Coarse exhaustive search*:** This search tests N_{srch} points in the domain of the parameter δ. Recalling (11) (for known CIR R_{l,k_p} is replaced by Y_{l,k_p}), it follows that evaluating the sequence $\left\{Y_{l,k_p}=Z_{l,k_p}D_{l,k_p}^{*}H_{k_p}^{*}\right\}_{l=1,p=1}^{L,K_p}$ requires LK_p complex multiplications.[§] Note that the terms $e^{-j\xi_l k_p\delta}$ can be evaluated in advance for the search points in δ. Thus, calculating the absolute value in (11), for a fixed l, requires K_p complex multiplications, K_p-1 complex additions, and one absolute value operation. This is done for each $l=1,2,\ldots,L$, and the summation over l requires $L-1$ real additions. Thus, the curly brackets in (11) require LK_p complex multiplications, $L(K_P-1)$ complex additions, L absolute value operations, and $L-1$ real additions. We evaluate N_{srch} values of δ. Thus, the course search requires $N_{\text{srch}}(3LK_p+L-1)$ flops.

***Cubic spline interpolation*:** Let N_{expan} be the interpolation expansion factor. From [33, Table 2, for Δ_4 and $m=3$] it follows that computing the interpolating polynomials requires $60+3N_{\text{srch}}$ real additions, $36+3N_{\text{srch}}$ real multiplications, and $18+2N_{\text{srch}}$ real divisions. These polynomials of degree 3 are evaluated at $N_{\text{srch}}(N_{\text{expan}}-1)$ values of δ. The evaluation of a polynomial of degree 3 requires three real multiplications and three real additions. The maximization requires $N_{\text{srch}}N_{\text{expan}}$ real comparisons. Thus, the spline interpolation requires $7N_{\text{srch}}N_{\text{expan}}+2N_{\text{srch}}+114$ flops.

***Best linear unbiased estimator based on the vector $\boldsymbol{\alpha}$*:** Recall $\hat{\alpha}_l\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)$ in Subsection 3.3. Evaluating the maximization argument in $\hat{\alpha}_l\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)$ requires carrying out K_p complex exponents, K_p complex multiplications, and K_p-1 complex additions. Furthermore, for each $l,l=1,2,\ldots,L$, an argument operation and a single real multiplication are required. Since the vector $\boldsymbol{\alpha}$ consists of L elements, it follows that evaluating $\boldsymbol{\alpha}$ requires LK_p complex multiplications, LK_p complex exponents, $L(K_p-1)$ complex additions, L argument operations, and L real multiplications. Next, let $\mathbf{b}=[N-1+2N_1,N-1+2N_2,\ldots,N-1+2N_L]$, and let $\boldsymbol{\nu}$ be a vector representing the errors in estimating $\alpha_l\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)$. We can represent the relationship between the estimates $\hat{\boldsymbol{\alpha}}\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)=\left[\hat{\alpha}_1\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right),\hat{\alpha}_2\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right),\ldots,\hat{\alpha}_L\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)\right]^T$ and ϵ via $\hat{\boldsymbol{\alpha}}\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)=\mathbf{b}\epsilon+\boldsymbol{\nu}$. Define $\gamma\triangleq\frac{N^2\text{F}_{N,L}}{\pi^2}\sum_{k_p\in\mathcal{K}_p}k_p^2|H_{k_p}|^2$ and define the vector $\mathbf{u}\triangleq\frac{N^2\sum_{k_p\in\mathcal{K}_p}k_p|H_{k_p}|^2}{\pi^2}\cdot[N-1+2N_1,N-1+2N_2,\ldots,N-1+2N_L)]^T$. Let the matrix $\mathbb{D}$ be an $L\times L$ diagonal matrix with the elements $[\mathbb{D}]_{l,l}=\frac{N^2}{\pi^2}\sum_{k_p\in\mathcal{K}_p}|H_{k_p}|^2, l=1,2,\ldots,L$. By following steps similar to those detailed in Appendix A, we have that the CRB for the estimation of $\boldsymbol{\alpha}$ is given by $\mathbb{C}=\left(\mathbb{D}-\frac{\mathbf{u}\mathbf{u}^T}{\gamma}\right)^{-1}$. Next, the BLUE for ϵ can be computed following [28, Chapter 6]: $\hat{\epsilon}_{\text{DML}}^{(\text{k})}=\left(\mathbf{b}^T\mathbb{C}\mathbf{b}\right)^{-1}\mathbf{b}^T\mathbb{C}\hat{\boldsymbol{\alpha}}\left(\hat{\delta}_{\text{DML}}^{(\text{k})}\right)$. It can be observed that $\mathbb{C}$ consists of: 1) Constants which can be calculated in advance. 2) The quantities $\left|H_{k_p}\right|^2, p=1,2,\ldots,K_p$, the terms $Q(q), q=1,2,3$, and the term $Q^2(2)$. These quantities can be evaluated using K_p complex multiplications, $2K_p+1$ real multiplications, and $3K_p-3$ real additions.

Given that $Q(q), q=1,2,3$ were evaluated, we have that evaluating the diagonal elements of $\mathbb{D}$ require a single real multiplication, evaluating $\frac{1}{\gamma}$ require a single real multiplication and a single real division, and evaluating $\mathbf{u}$ requires $L+1$ real multiplications. Each diagonal element of $\mathbb{C}$ is evaluated using a real addition and a real multiplication, while the off-diagonal elements require a single multiplication. Thus, evaluating $\mathbb{C}$ requires L^2+L+3 real multiplications, L real additions, and a single real division. Next, evaluating $\mathbf{b}^T\mathbb{C}$ requires L^2 real multiplications and $L(L-1)$ real additions. Then, evaluating $\mathbf{b}^T\mathbb{C}\mathbf{b}$ requires L real multiplications and $L-1$ real additions. The explicit evaluation of $\hat{\epsilon}_{\text{DML}}^{(\text{k})}$ requires L real multiplications and $L-1$ real additions. Finally, a single real division operation is required. Thus, the BLUE requires $3L^2+3LK_p+7L+5K_p+1$ flops.

We conclude that the DML requires the following number of operations (complex and real):

$$\mathscr{C}_{\text{DML}}=3N_{\text{srch}}LK_p+7N_{\text{srch}}N_{\text{expan}}+N_{\text{srch}}L+3L^2+3LK_p+N_{\text{srch}}+7L+5K_p+115.$$

D.2. Computational complexity analysis for the AOPSE of Subsection 4.2

Recall that the AOPSE method consists of two steps: 1) $\hat{\rho}_{k_p,\text{poly}}^{(\text{k})}$ is estimated via the polynomial approximation for

[§] As the symbols at the pilots subcarriers are selected from a BPSK constellation, calculating $D_{l,k_p}^{*}H_{k_p}^{*}$ does not require a multiplication operation.

each subcarrier, k_p (15). 2) $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ is estimated out of $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}$ via the BLUE (18).

Evaluating the coefficients of the approximating polynomials of degree M: The terms $\xi_l, l = 1,2,\ldots,L$ can be calculated in advance. Evaluating the coefficients requires $2(M+1)L$ real multiplications and $2(M+1)L-1$ real additions.

Finding the desired root via Newton-Raphson per subcarrier: The NR method, see [30], successively finds better approximations to the roots of a real-valued function. Let $f(x), x \in \Re$, be a real-valued function, and let $f'(x)$ denote its first derivative. Let x_0 be an initial guess for the root of $f(x)$. Then, an approximation of the a root of $f(x)$ can be successively obtained via $x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, n = 1,2,\ldots,N_{\mathrm{NR}}$, where N_{NR} is the number of successive iterations. In our work, the function $f(x)$ is a polynomial of degree $2M+1$ with real coefficients, and therefore $f'(x)$ is a polynomial of degree $2M$ with real coefficients.

We begin with evaluating the coefficients of $f'(x)$. This requires $2M$ real multiplications. At the first iteration, that is, $n = 1$, as x_0 is known, the powers $x_0^m, m = 2,3,\ldots,2M+1$ are known as well. Therefore, evaluating $\frac{f(x_0)}{f'(x_0)}$ requires $4M-1$ real multiplications, $4M$ real additions, and a single real division. For $n = 2,3,\ldots,N_{\mathrm{NR}}$, evaluating $f(x_{n-1})$ requires $4M-1$ real multiplications ($2M-1$ for evaluating the powers $\{x_{n-1}^m\}_{m=1}^{2M+1}$, and $2M$ for the multiplication of the polynomial coefficients and the evaluated powers), and $2M$ real additions. Evaluating $f'(x_{n-1})$ requires $2M-2$ real multiplications, $2M-1$ real additions, and evaluating $\frac{f(x_{n-1})}{f'(x_{n-1})}$ requires a single real division. Recall that the NR method is applied in our algorithm for N_0 different initial conditions. Finally, the desired root is selected by evaluating (14), which requires $N_0 L$ complex exponents, $N_0 L$ complex multiplications, $N_0(L-1)$ real additions, and N_0 real comparisons.

Thus, finding the desired root via NR, for K_p subcarriers, requires $K_p(2M + N_0(10N_{\mathrm{NR}}M - 3N_{\mathrm{NR}} - 2M + 3L + 3))$ flops.

Estimation of $\boldsymbol{\theta}$ via the best linear unbiased estimator: Recall the definition of $\mathbb{A}$: $\mathbb{A}$ is a $K_p \times 2$ matrix whose elements are $[\mathbb{A}]_{p,1} = 1, [\mathbb{A}]_{p,2} = k_p, p \in \{1,2,\ldots,K_p\}$. Let $\boldsymbol{\nu}$ be a vector representing the errors in estimating $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}$. We can represent the relationship between the estimates $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}$ and $\boldsymbol{\theta}$ via $\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}} = \mathbb{A}\boldsymbol{\theta} + \boldsymbol{\nu}$, where the covariance matrix of $\boldsymbol{\nu}$, $\mathbb{Q}_\nu$, is a diagonal matrix with diagonal entries which are given by $[\mathbb{Q}_\nu]_{p,p} = \frac{N^2\sigma_v^2}{2\pi^2\left|H_{k_p}\right|^2 \mathrm{F}_{N,L}}$. Now, $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ is given by $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} = \left(\mathbb{A}^H\mathbb{Q}_\nu^{-1}\mathbb{A}\right)^{-1}\mathbb{A}^H\mathbb{Q}_\nu^{-1}\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}}$. Plugging $\mathbb{A}$ and $\mathbb{Q}_\nu$ we have $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ in (D.1) at the bottom of the page. Now, the p'th column of the matrix $\mathbb{D}$, in (D.1), is given by:

$$\begin{bmatrix}[\mathbb{D}]_{1,p}\\ [\mathbb{D}]_{2,p}\end{bmatrix} = \begin{bmatrix} Q^\star(3)\left|H_{k_p}\right|^2 - Q^\star(2)\left|H_{k_p}\right|^2 k_p \\ Q^\star(1)\left|H_{k_p}\right|^2 k_p - Q^\star(2)\left|H_{k_p}\right|^2\end{bmatrix}.$$

Evaluating $\left|H_{k_p}\right|^2, p = 1,2,\ldots,K_p$, requires K_p complex multiplications. Evaluating $Q^\star(q), q = 1,2,3$, requires $2K_p+5$ real multiplications, $3K_p-2$ real additions, and a single real division. The matrix $\mathbb{D}$ is evaluated using $4K_p$ real multiplications and $2K_p$ real additions. Finally, evaluating $\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}}$ requires $2K_p$ real multiplications and $2K_p-2$ real additions. Thus, estimation of $\boldsymbol{\theta}$ via the BLUE requires $16K_p+2$ flops.

We conclude that the AOPSE requires the following number of operations (complex and real):

$$\begin{aligned}\mathscr{C}_{\mathrm{AOPSE}} &= K_p(4L(M+1) + 2M + 15\\ &\quad + N_0(10N_{\mathrm{NR}}M - 3N_{\mathrm{NR}} - 2M + 3L + 3)) + 2.\end{aligned}$$

ACKNOWLEDGEMENTS

This work was supported by the Israeli Ministry of Economy through the Israeli Smart Grid Consortium. Parts of this work were presented at the IEEE Wireless Communications and Networking Conference (WCNC), April 2014, Istanbul, Turkey.

REFERENCES

1. Digital video broadcasting (DVB): frame structure, channel coding and modulation for digital terrestrial television. ETSI EN 300 744, 2001.
2. IEEE standard for air interface for broadband wireless access systems. IEEE Std. 802.16, 2012.
3. Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation. 3GPP TS 36.211, 2012.
4. Steendam H, Moeneclaey M. Synchronization sensitivity of multicarrier systems. *European Transactions on Telecommunications* 2004; **15**(3): 223–234.
5. Pollet T, Van Bladel M, Moeneclaey M. "BER sensitivity of OFDM systems to carrier frequency offset and Wiener phase noise," *IEEE Transactions on Communications* 1995; **43**(2/3/4): 191–193.
6. Schmidl TM, Cox DC. Robust frequency and timing synchronization for OFDM. *IEEE Transactions on Communications* 1997; **45**(12): 1613–1621.

$$\hat{\boldsymbol{\theta}}^{(\mathrm{k})}_{\mathrm{poly}} = \underbrace{\begin{bmatrix} Q^\star(3) & -Q^\star(2) \\ -Q^\star(2) & Q^\star(1)\end{bmatrix}\cdot\begin{bmatrix} \left|H_{k_1}\right|^2 & \left|H_{k_2}\right|^2 & \ldots & \left|H_{k_{K_P}}\right|^2 \\ \left|H_{k_1}\right|^2 k_1 & \left|H_{k_2}\right|^2 k_2 & \ldots & \left|H_{k_{K_P}}\right|^2 k_{K_P}\end{bmatrix}}_{\mathbb{D}}\cdot\hat{\boldsymbol{\rho}}^{(\mathrm{k})}_{\mathrm{poly}} \tag{D.1}$$

7. Morelli M, Mengali U. An improved frequency offset estimator for OFDM applications. *IEEE Communications Letters* 1999; **3**(3): 75–77.
8. Cvetkovic Z, Tarokh V, Yoon S. On frequency offset estimation for OFDM. *IEEE Transactions on Wireless Communications* 2013; **12**(3): 1062–1072.
9. Ziabari HA, Shayesteh MG. Robust timing and frequency synchronization for OFDM Systems. *IEEE Transactions on Vehicular Technology* 2011; **60** (8): 3646–3656.
10. Dai L, Wang Z, Wang J, Yang Z. Joint channel estimation and time-frequency synchronization for uplink TDS-OFDMA systems. *IEEE Transactions on Consumer Electronics* 2010; **56**(2): 494–500.
11. Dai L, Zhang C, Xu Z, Wang Z. Spectrum-efficient coherent optical OFDM for transport networks. *IEEE Journal of Selected Areas in Communications* 2010; **31**(1): 62–74.
12. Speth M, Fechtel SA, Fock G, Meyr H. Optimum receiver design for wireless broad-band systems using OFDM – part I. *IEEE Transactions on Communications* 1999; **47**(11): 1668–1677.
13. Nguyen-Le H, Le-Ngoc T, Ko CC. RLS-based joint estimation and tracking of channel response, sampling, and carrier frequency offsets for OFDM. *IEEE Transactions on Broadcasting* 2009; **55**(1): 84–94.
14. Jose R, Ambat SK, Hari KVS. Low complexity joint estimation of synchronization impairments in sparse channel for MIMO-OFDM system. *International Journal of Electronics and Communications* 2014; **68**(2): 151–157.
15. Weikert O. Joint estimation of carrier and sampling frequency offset, phase noise, IQ offset and MIMO channel for LTE Advanced UL MIMO. In *Proceedings of IEEE 14th Workshop on Sig. Proc. Advance in Wireless Communications*, Darmstadt, Germany, 2013; 520–524.
16. Laourine A, Stéphenne A, Affes S. Joint carrier and sampling frequency offset estimation based on harmonic retrieval. In *Proceedings of Vehicular Technology Conference (VTC 2006 Fall)*, Montreal, Canada, 2006; 1–5.
17. Häring L, Bieder S, Czylwik A. "Residual carrier and sampling frequency synchronization in multiuser OFDM systems. In *Proceedings of Vehicular Technology Conference (VTC 2006 Spring)*, Melbourne, Australia, 2006; 1937–1941.
18. Oberli C. ML-based tracking algorithms for MIMO-OFDM. *IEEE Transactions on Wireless Communications* 2007; **6**(7): 2630–2639.
19. Guo K, Xu W, Zhou G. Differential carrier frequency offset and sampling frequency offset estimation for 3GPP LTE. In *Proceedings of Vehicular Technology Conference (VTC 2011 Spring)*, Yokohama, Japan, 2011; 1–5.
20. Speth M, Fechtel SA, Fock G, Meyr H. Optimum receiver design for OFDM-based broadband transmission – part II: A case study. *IEEE Transactions on Communications* 2001; **49**(4): 571–578.
21. Shin W, Lee S, You Y. A robust joint frequency offset estimation for the OFDM system using cyclic delay diversity. *Wireless Personal Communications* 2014; **77**(4): 2483–2496.
22. Yang H, Shin W, Lee S, You Y. A robust estimation of residual carrier frequency offset with I/Q imbalance in OFDM systems. *IEEE Transactions on Vehicular Technology* 2014. **DOI:** 10.1109/TVT.2014.2311157
23. Liu SY, Chong JW. A study of joint tracking algorithms of carrier frequency offset and sampling clock offset for OFDM-based WLANs. In *Proceedings of IEEE International Conference in Communications, Circuits and Systems and West Sino Expositions*, Chengdu, China, 2002; 109–113.
24. Sun Z, Yu L, Ren C. Joint estimation of carrier and sampling frequency offset for OFDM systems in slow fading channel. In *Procedings of IEEE International Conference in Communications and Technology*, Chengdu, China, 2012; 148–154.
25. Morelli M, Moretti M. Fine carrier and sampling frequency synchronization in OFDM systems. *IEEE Transactions on Wireless Communications* 2010; **9**(4): 1514–1524.
26. Kim Y, Lee J. Joint maximum likelihood estimation of carrier and sampling frequency offsets for OFDM systems. *IEEE Transactions on Broadcasting* 2011; **57**(2): 277–283.
27. Wang X, Hu B. A low-complexity ML estimator for carrier and sampling frequency offsets in OFDM systems. *IEEE Communications Letters* 2014; **18**(3): 503–506.
28. Kay SM. *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall: Englewood Cliffs, NJ, 1993.
29. Evolved universal terrestrial radio access (E-UTRA); physical layer procedures. 3GPP TS 36.213, 2012.
30. Mekwi WR. Iterative methods for roots of polynomials. *Master's Thesis*, University of Oxford, 2001.
31. Rockah Y, Schultheiss PM. Array shape calibration using sources in unknown locations-part I: far-field sources. *IEEE Transactions on Acoustics, Speech and Signal Processing* 1987; **35**(3): 286–299.
32. Mengali U, Morelli M. Data-aided frequency estimation for burst digital transmission. *IEEE Transactions on Communications* 1997; **45**(1): 23–25.

33. Toraichi K, Katagishi K, Sekita I, Mori R. Computational complexity of spline interpolation. *International Journal of Systems Science* 1987; **18** (5): 945–954.
34. Boyd S, Vandenberghe L. *Convex Optimization*. Cambridge University Press: Cambridge, UK, 2004.

AUTHORS' BIOGRAPHIES

Yonathan Murin received his BSc degree in Electrical Engineering and Computer Science from Tel-Aviv University, Israel, in 2004, and his MSc (magna cum laude) degree in Electrical Engineering from Ben-Gurion University, Israel, in 2011. He is currently pursuing a PhD degree in Electrical Engineering at Ben-Gurion University. From 2004 to 2010, he worked as a DSP and algorithms engineer, and as a team leader at Comsys Communication and Signal Processing. His research interests include network information theory and wireless communications.

Ron Dabora received his BSc and MSc degrees in 1994 and 2000, respectively, from Tel-Aviv University, and his PhD degree in 2007 from Cornell University, all in Electrical Engineering. From 1994 to 2000, he worked as an engineer at the Ministry of Defense of Israel, and from 2000 to 2003, he was with the Algorithms Group at Millimetrix Broadband Networks, Israel. From 2007 to 2009, he was a postdoctoral researcher at the Department of Electrical Engineering, Stanford University. Since 2009, he is an assistant professor at the Department of Electrical and Computer Engineering, Ben-Gurion University, Israel. His research interests include network information theory, wireless communications, and power line communications. Dr Dabora served as a TPC member in a number of international conferences including WCNC, PIMRC, and ICC. From 2012 to 2014, he served as an associate editor for the *IEEE Signal Processing Letters*, and he currently serves as a senior area editor for the *IEEE Signal Processing Letters*.