Joint Carrier Frequency Offset and Channel Impulse Response Estimation for Linear Periodic Channels

Roee Shaked, Nir Shlezinger, and Ron Dabora

Abstract—We study joint estimation of the channel impulse response (CIR) and of the carrier frequency offset (CFO) for linear channels in which both the CIR and the noise statistics vary periodically in time. This model corresponds to interference-limited communications as well as to power line communication and doubly selective channels. We first consider the joint maximum likelihood estimator (JMLE) for the CIR and the CFO and show it has a high computational complexity and relatively low spectral efficiency. This motivates the derivation of two estimation schemes with higher spectral efficiency and lower computational complexity compared to the JMLE, obtained by exploiting both the periodicity of the channel and the fact that, typically, the delay–Doppler spreading function of the CIR is approximately sparse, *without requiring a-priori knowledge of the sparsity pattern*. The proposed estimation schemes are numerically tested and the results demonstrate that substantial benefits can be obtained by properly accounting for the approximate sparsity and periodicity in the design of the estimation scheme.

I. Introduction

Many practical communications channels exhibit periodic characteristics. An important example is interference-limited communications: As man-made signals typically obey a periodic statistical model [1, Sec. 5], the appropriate channel model for interference-limited communications scenarios, arising, e.g., in 5G cellular [2], exhibits periodic characteristics. Periodicity is also inherent to narrowband (NB) power line communications (PLC) [3], due to the periodicity of the mains voltage across the grid. Lastly, we note that doubly selective channels can also be modeled as systems having periodically varying channel impulse response (CIR) [4]. It follows that studying communications over liner periodic channels is very relevant for future communications scenarios.

Generally speaking, high data rate communications requires accurate knowledge of the CIR and of the carrier frequency offset (CFO), caused by instabilities of the oscillators [5, Sec. 1]. This need has motivated the proposal of many training–based schemes for estimating these parameters using an a-priori known pilot sequence (PS), while modeling the CIR and the CFO as *unknown deterministic parameters* [5, Sec. 3.1]. The work [6] designed a maximum likelihood (ML) scheme for jointly estimating the CIR and the CFO for linear time-invariant (LTI) channels with additive white Gaussian noise (AWGN). Joint ML estimation of the CIR and the CFO for orthogonal frequency division multiplexing (OFDM) modulated signals was studied in [5, Sec. 3.1.1]. In [7] joint estimation was studied for point-to-point multiple input-multiple output (MIMO) scenarios, and in [5, Sec. 4] it was studied for MIMO-OFDM scenarios. The work [8] proposed a CFO estimation scheme for periodic channels *without estimating the unknown CIR coefficients*. Yet, none of these works considered joint CFO and CIR estimation over linear periodic channels, modeled as having linear periodically time-varying (LPTV) CIRs with additive cyclostationary Gaussian noise (ACGN).

As evident from [4] and [7], reducing the number of CIR coefficients to be estimated improves estimation performance. A common approach to reducing the number of estimated parameters is to assume that the CIR is sparse, and employ compressed sensing (CS) techniques [9], [10]. Indeed, in [4], it was shown that applying CS improves the performance of CIR estimation *for wireless channels* as well as the spectral efficiency. The works [11], [12] proposed a CS-based joint CIR and CFO estimator for slow-fading wireless LTI channels with AWGN. However, in [11] it was assumed that the number of non-zero CIR coefficients is *a-priori known*, and in [12] it was assumed that the number of observations *is larger* than the number of estimated parameters. Lastly, the work [13] designed an estimator for LPTV CIRs with *temporally uncorrelated* ACGN, assuming that the noise variance, the CFO, and the sparsity pattern of the CIR coefficients are all *a-priori known*. To the best of our knowledge, joint estimation of LPTV CIR coefficients and of CFO in the presence of correlated ACGN, with or without the use of CS techniques, has not yet been studied.

Main Contributions: In this work we study joint CIR and CFO for LPTV channels with ACGN. We first present the joint ML estimator (JMLE) and explicitly characterize its drawbacks, namely, high computational complexity and low spectral efficiency. Then, to improve upon these drawbacks, we research the application of sparsity–based compression techniques. We first conclude that *when the CFO is unknown, then the sparsity pattern is also unknown*. This motivates us to propose two JMLE-oriented schemes utilizing the approximate sparsity of the CIR while not assuming a-priori knowledge of the sparsity pattern. Note that the application of CS decreases the number of parameters needed for representing the CIR, hence decreasing the length of the PS and improving spectral efficiency. We demonstrate via a simulation study that the proposed schemes outperform the JMLE over a wide range of signal-to-noise ratios (SNRs), while increasing the spectral efficiency and reducing the computational complexity.

The rest of this paper is organized as follows: Section II formulates the problem and details the JMLE of the CFO and the CIR; Section III derives two JMLE-oriented estimators which utilize compression techniques; Section IV details the results of a numerical study; Lastly, Section V provides some concluding remarks.

II. Preliminaries

Notations: We denote the sets of non-negative integers, integers, and complex numbers by $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{C}$, respectively.

R. Shaked, N. Shlezinger, and R. Dabora are with the department of Electrical and Computer Engineering, Ben-Gurion University, Beer-Sheva, Israel (e-mail: {shroee, nirshl}@post.bgu.ac.il; ron@ee.bgu.ac.il).

$|\mathcal{X}|$ denotes the cardinality of the set $\mathcal{X}$. Column vectors are denoted with lower-case boldface letters; and the k-th element $(k \geq 0)$ of a vector $\mathbf{x}$ is denoted with $[\mathbf{x}]_k$. Matrices are denoted with upper-case boldface letters, where the element at the k-th row and the l-th column of a matrix $\mathbf{X}$ is denoted by $[\mathbf{X}]_{k,l}$. $\mathbf{I}_M$ denotes the $M \times M$ identity matrix. The Kronecker product, Hermitian transpose, transpose, complex conjugate, Euclidean norm, and stochastic expectation are denoted by $\otimes$, $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^*$, $\|\cdot\|$, and $\mathbb{E}\{\cdot\}$, respectively. $\langle x \rangle$ denotes the fractional part of a real number x s.t. $\langle x \rangle \in (-1/2, 1/2] \triangleq \mathcal{I}_{1/2}$, $\lfloor x \rfloor$ denotes the largest integer not greater than x, and for $x \in \mathbb{Z}$, $((x))_m$ denotes the remainder of x when divided by $m \in \mathbb{N}$. $\mathcal{CN}$ denotes the proper-complex (PC) Gaussian distribution.

Problem Formulation: We consider a discrete-time, scalar, finite-memory, periodic channel, modeled as an LPTV filter with ACGN. Let $w[n]$ denote the PC zero-mean ACGN with autocorrelation function $c_w[n,l] \triangleq \mathbb{E}\{w[n]w^*[n+l]\}$, and let $L_{\text{noise}}, N_{\text{noise}} \in \mathbb{N}$ denote *the length and the period of the temporal correlation of the noise*, respectively. Thus, $c_w[n,l] = c_w[n+N_{\text{noise}},l]$, $\forall n, l \in \mathbb{Z}$ (see also [1, Sec. 3.5]), and $c_w[n,l] = 0$ for all $|l| \geq L_{\text{noise}}$. Next, let $g[n,l]$ denote the causal LPTV CIR, let $L_{\text{ch}}, N_{\text{ch}} \in \mathbb{N}$ denote *the length and the period of the CIR*, respectively, i.e., $g[n,l] = g[n+N_{\text{ch}},l]$, $\forall n, l \in \mathbb{Z}$, and $g[n,l] \neq 0$ only for $0 \leq l < L_{\text{ch}}$. The CIR is assumed to be deterministic and unknown to the receiver. Set N_0 to be the least common multiple of N_{ch} and N_{noise}. We refer to N_0 as the *period of the channel*. Let $\nu \in \mathcal{I}_{1/2}$ denote the CFO between the transmitter and the receiver, and let ϕ_0 denote the phase offset at time $n = 0$. Lastly, let $x[n]$ and $y[n]$ to denote the channel input and the channel output, respectively, which are related via:

$$y[n] = e^{j(2\pi\nu n + \phi_0)} \sum_{l=0}^{L_{\text{ch}}-1} x[n-l]g[n,l] + w[n], \qquad n \in \mathbb{Z}. \quad (1)$$

Define the parameter $\alpha \triangleq \langle N_0 \cdot \nu \rangle$, which measures the phase offset accumulated during a single period, and let

$$h[n,l] \triangleq e^{j(2\pi\nu((n))_{N_0} + \phi_0)} g[n,l], \quad (2)$$

represent the phase–shifted CIR (recall that the phase shift is unknown). Note that $h[n,l]$ inherits the periodicity and finite memory of $g[n,l]$, Plugging (2) into (1) yields

$$y[n] = \sum_{l=0}^{L_{\text{ch}}-1} x[n-l] e^{j2\pi\alpha\lfloor \frac{n}{N_0} \rfloor} h[n,l] + w[n]. \quad (3)$$

Since $h[n,l]$ is periodic, it has a discrete Fourier series (DFS) [14, Ch. 8.1] with DFS coefficients $\underline{h}[k,l]$. $k \in \{0,1,...,N_0-1\} \triangleq \mathcal{N}_0$, $l \in \{0,1,...,L_{\text{ch}}-1\} \triangleq \mathcal{L}_{\text{ch}}$. We refer to $\underline{h}[k,l]$ as the delay-Doppler spreading function (DDSF), as in [4]. Note that $h[n,l]$ can be *uniquely recovered* from $\underline{h}[k,l]$ via the inverse DFS (IDFS). Expressing $h[n,l]$ in (3) using its DFS yields

$$y[n] = \sum_{l=0}^{L_{\text{ch}}-1} \frac{x[n-l]}{\sqrt{N_0}} \sum_{k=0}^{N_0-1} e^{j\frac{2\pi}{N_0}nk} e^{j2\pi\alpha\lfloor \frac{n}{N_0} \rfloor} \underline{h}[k,l] + w[n]. \quad (4)$$

In the following we design a scheme for estimating the unknown parameters $(\alpha, \{\underline{h}[k,l]\}_{k \in \mathcal{N}_0, l \in \mathcal{L}_{\text{ch}}})$, assuming that the receiver knows the statistical moments of the noise.

The JMLE: Let $N_{\text{CIR}} \triangleq N_0 L_{\text{ch}}$ denote the overall number of CIR coefficients and let N_{obs} denote the number of observations used for estimation. For simplicity, we assume that N_{obs}

is an integer multiple of the channel period, $N_{\text{obs}} = N_{\text{p}} N_0$, $N_{\text{p}} \in \mathbb{N}$. The channel input consists of a predefined PS of length $N_{\text{seq}} = N_{\text{obs}} + L_{\text{ch}} - 1$, denoted as $\{s[n]\}_{n=-L_{\text{ch}}+1}^{N_{\text{obs}}-1}$. When the PS is transmitted, the channel outputs at time indexes $n \in \{0,1,..,N_{\text{obs}}-1\} \triangleq \mathcal{N}_{\text{obs}}$ are obtained from (4) by setting $x[n] = s[n]$. Next, define the vectors $\mathbf{y}, \mathbf{w} \in \mathbb{C}^{N_{\text{obs}}}$ s.t. $[\mathbf{y}]_n = y[n]$, $[\mathbf{w}]_n = w[n]$, $n \in \mathcal{N}_{\text{obs}}$, and define the $N_{\text{obs}} \times N_{\text{obs}}$ matrix $\mathbf{C_w} \triangleq \mathbb{E}\{\mathbf{ww}^H\}$ s.t. $[\mathbf{C_w}]_{n_1,n_2} = c_w[n_1, n_2 - n_1]$. We also define the DDSF coefficients vector $\underline{\mathbf{h}} \in \mathbb{C}^{N_{\text{CIR}}}$ s.t. $[\underline{\mathbf{h}}]_{kL_{\text{ch}}+l} = \underline{h}[k,l]$, $k \in \mathcal{N}_0$, $l \in \mathcal{L}_{\text{ch}}$, and the PS matrix $\mathbf{S} \in \mathbb{C}^{N_{\text{obs}} \times N_{\text{p}} N_{\text{CIR}}}$ s.t. $[\mathbf{S}]_{n,l} = s[n - ((l))_{L_{\text{ch}}}]$ for $n \in \mathcal{N}_{\text{obs}}$, $nL_{\text{ch}} \leq l < (n+1)L_{\text{ch}}$ and $[\mathbf{S}]_{n,l} = 0$ otherwise. Let $\mathbf{F}_{N_0}^{-1}$ be the $N_0 \times N_0$ IDFS matrix, i.e., $[\mathbf{F}_{N_0}^{-1}]_{n,k} = \frac{1}{\sqrt{N_0}} e^{j\frac{2\pi}{N_0}kn}$ and define the matrix $\mathbf{F} \in \mathbb{C}^{N_{\text{p}} N_{\text{CIR}} \times N_{\text{p}} N_{\text{CIR}}}$ s.t. $\mathbf{F} \triangleq \mathbf{I}_{N_{\text{p}}} \otimes \mathbf{F}_{N_0}^{-1} \otimes \mathbf{I}_{L_{\text{ch}}}$. Lastly, define the matrix $\mathbf{\Phi}_{(\alpha)} \in \mathbb{C}^{N_{\text{p}} N_{\text{CIR}} \times N_{\text{CIR}}}$ s.t. $\mathbf{\Phi}_{(\alpha)} \triangleq [1, e^{j2\pi\alpha}, ..., e^{j2\pi(N_{\text{p}}-1)\alpha}]^T \otimes \mathbf{I}_{N_{\text{CIR}}}$, as well as the matrix $\mathbf{H}_{(\alpha)} \triangleq \mathbf{SF\Phi}_{(\alpha)}$, and the matrix $\mathbf{G}_{(\alpha)} \triangleq (\mathbf{H}_{(\alpha)}^H \mathbf{C_w}^{-1} \mathbf{H}_{(\alpha)})^{-1}$. Applying these definitions to (4), we arrive at:

$$\mathbf{y} = \mathbf{H}_{(\alpha)} \underline{\mathbf{h}} + \mathbf{w}, \qquad \mathbf{y} \sim \mathcal{CN}(\mathbf{H}_{(\alpha)} \underline{\mathbf{h}}, \mathbf{C_w}). \quad (5)$$

When the number of observations is *larger* than the number of estimated parameters, i.e., $N_{\text{p}} > L_{\text{ch}}$, and $\mathbf{H}_{(\alpha)}$ has a full rank, then the JMLE of α and $\underline{\mathbf{h}}$ based on $\mathbf{y}$, denoted $(\hat{\alpha}^{\text{JML}}, \hat{\underline{\mathbf{h}}}^{\text{JML}})$, can be obtained from (5) following steps as in [15, Ch. 8.9]:

$$\hat{\alpha}^{\text{JML}} = \underset{\alpha \in \mathcal{I}_{1/2}}{\arg\max} \left\{ \mathbf{y}^H \mathbf{C_w}^{-1} \mathbf{H}_{(\alpha)} \mathbf{G}_{(\alpha)} \mathbf{H}_{(\alpha)}^H \mathbf{C_w}^{-1} \mathbf{y} \right\}, \quad (6a)$$

and

$$\hat{\underline{\mathbf{h}}}_{(\hat{\alpha}^{\text{JML}})}^{\text{JML}} = \mathbf{G}_{(\hat{\alpha}^{\text{JML}})} \mathbf{H}_{(\hat{\alpha}^{\text{JML}})}^H \mathbf{C_w}^{-1} \mathbf{y}. \quad (6b)$$

Note that when $g[n,l]$ is LTI, and the noise is AWGN, then, the JMLE in (6) coincides with the JMLE proposed in [6].

Note that the JMLE has two major drawbacks: First, the estimate $\hat{\alpha}^{\text{JML}}$ is obtained via a grid search which requires evaluating (6a) at T test points. Assuming that $T >> N_{\text{CIR}}, N_{\text{obs}}$ the complexity of calculating the estimate (6) is dominated by T inversions of an $N_{\text{CIR}} \times N_{\text{CIR}}$ complex matrix $\mathbf{G}_{(\alpha)}$, and thus, the computational complexity of the JMLE is on the order of $\mathcal{O}(T \cdot N_{\text{CIR}}^3)$. Moreover, to facilitate the estimation, N_{p} must be larger than L_{ch}, resulting in a low spectral efficiency. We next propose two modifications of (6) which improve the spectral efficiency and reduce the computational complexity, by utilizing the approximate sparsity of the DDSF coefficients.

III. ESTIMATION SCHEMES BASED ON UTILIZING SPARSITY

Assuming that $\underline{\mathbf{h}}$ can be *approximated as a sparse vector*, see, e.g., [4], [13], then there exists a set of indexes $\mathcal{K} \subseteq \{0, 1, \ldots, N_{\text{CIR}} - 1\} \triangleq \mathcal{N}_{\text{CIR}}$, s.t. $[\underline{\mathbf{h}}]_k \approx 0$ for each $k \notin \mathcal{K}$. The set $\mathcal{K}$ is henceforth referred to as the *sparsity pattern*. Note that $\mathcal{K}$ represents both sparsity in the lag-domain, i.e., w.r.t $l \in \mathcal{L}_{\text{ch}}$, and in the Doppler-domain, i.e., w.r.t $k \in \mathcal{N}_0$. Recalling that $\underline{\mathbf{h}}$ is defined as the stacking of $\underline{h}[k,l]$, and letting $\underline{g}[k,l]$ denote the DDSF of the CIR in the absence of CFO and phase offset. The first result is a characterization of the relationship between the non-zero elements of $\underline{h}[k,l]$ and of $\underline{g}[k,l]$. This relationship is stated in the following lemma:

Lemma 1. $\underline{h}[k,l]$ *is equal to zero if and only if*

$$\sum_{k'=0}^{N_0-1} \underline{g}[k',l] \frac{\sin\left(\pi(k-k'-N_0\nu)\right)}{\sin\left(\frac{\pi(k-k'-N_0\nu)}{N_0}\right)} e^{-j\pi\frac{(N_0-1)}{N_0}(k-k'-N_0\nu)} = 0.$$

Proof: Denote $v[n] \triangleq e^{j(2\pi\nu((n))_{N_0}+\phi_0)}$. $v[n]$ is periodic with a period of N_0 and its DFS, denoted $\underline{v}[k]$, is given by

$$\underline{v}[k] = \frac{e^{j\phi_0}}{\sqrt{N_0}} \frac{\sin\left(\pi(k-N_0\nu)\right)}{\sin\left(\frac{\pi(k-N_0\nu)}{N_0}\right)} e^{-j\pi\frac{(N_0-1)}{N_0}(k-N_0\nu)}, \ k \in \mathcal{N}_0.$$ From

Eq. (2), $h[n,l] = g[n,l]v[n]$, $\forall l \in \mathcal{L}_{ch}$, hence, $\underline{h}[k,l]$ can be obtained as the circular convolution of length N_0 between $g[k,l]$ and $\underline{v}[k]$, see [14, Table 8.2]. Equating the circular convolution to zero proves the lemma. ∎

Lemma 1 implies that the *sparsity pattern $\mathcal{K}$ depends on ν*, hence, when ν is unknown, *the set $\mathcal{K}$ cannot be a-priori known,* and it has to be estimated. It follows that, if $\mathcal{K}$ is assumed a-priori known, then only the CIR may be assumed unknown, and joint estimation is irrelevant. Consequently, any joint estimation scheme based on sparsity, must incorporate a dedicated step for estimating $\mathcal{K}$. In the following two subsection we describe two such schemes.

A. Approximated ML with Sparsity Pattern Estimation

Existing sparsity estimation algorithms, e.g., [9], [10], are based on the assumption that the observations are modeled as a product between a *known* matrix and an unknown vector, possibly corrupted by an additive noise. To express the observations $\mathbf{y}$ in such a form, we define the vector $\underline{\mathbf{h}}_{\Phi} \in \mathbb{C}^{N_P N_{CIR}}$ s.t. $\underline{\mathbf{h}}_{\Phi} \triangleq \Phi_{(\alpha)}\underline{\mathbf{h}}$. Thus, (5) can be written as

$$\mathbf{y} = \mathbf{SF}\underline{\mathbf{h}}_{\Phi} + \mathbf{w}, \tag{7}$$

Letting $\mathcal{N}_P \triangleq \{0, 1, \ldots, N_P-1\}$ and recalling that $\mathcal{K}$ denotes the sparsity pattern of $\underline{\mathbf{h}}$, we next characterize the relationship between the sparsity patterns of $\underline{\mathbf{h}}_{\Phi}$ and of $\underline{\mathbf{h}}$:

Lemma 2. *The sparsity pattern of $\underline{\mathbf{h}}_{\Phi}$ is given by*

$$\underline{\mathcal{K}} = \left\{\underline{k} \in \mathbb{N} | \underline{k} = pN_{CIR} + k, p \in \mathcal{N}_P, k \in \mathcal{K}\right\}. \tag{8}$$

Proof: Since $\underline{\mathbf{h}}_{\Phi} \triangleq \Phi_{(\alpha)}\underline{\mathbf{h}}$, then, $[\underline{\mathbf{h}}_{\Phi}]_{pN_{CIR}+k} = e^{j2\pi\alpha p}[\underline{\mathbf{h}}]_k$, $\forall p \in \mathcal{N}_P$, $k \in \mathcal{N}_{CIR}$, and $\alpha \in \mathcal{I}_{1/2}$. Thus, $|[\underline{\mathbf{h}}_{\Phi}]_{pN_{CIR}+k}| = |e^{j2\pi p\alpha}[\underline{\mathbf{h}}]_k| = |[\underline{\mathbf{h}}]_k|$. It follows that $[\underline{\mathbf{h}}_{\Phi}]_{\underline{k}} = 0$ if and only if $[\underline{\mathbf{h}}]_{((\underline{k}))_{N_{CIR}}} = 0$, i.e., $\underline{k} \in \underline{\mathcal{K}}$ if and only if $((\underline{k}))_{N_{CIR}} \in \mathcal{K}$. ∎

It follows from Lemma 2 that the sparsity pattern of $\underline{\mathbf{h}}_{\Phi}$ consists of the sparsity pattern of $\underline{\mathbf{h}}$ replicated N_P times, therefore, $\underline{\mathbf{h}}_{\Phi}$ exhibits *group sparsity* [10], and $\mathcal{K} = \underline{\mathcal{K}} \cap \mathcal{N}_{CIR}$. Leveraging the group sparsity and the relationship (7), we propose to estimate the sparsity pattern using the adaptation of the basis pursuit denoising (BPDN) algorithm [9] to signals possessing structured sparsity patterns, as detailed in [10]. The proposed sparsity estimation scheme consists of three steps:

1) Fix a positive σ^2 such that $\sigma^2 \geq \mathbb{E}\{\|\mathbf{w}\|^2\}$.
2) Use the BPDN-based scheme detailed in [10] to recover the group sparse vector $\underline{\mathbf{h}}_{\Phi}$. This is achieved by solving the convex minimization problem:

$$\underline{\mathbf{h}}_{\Phi}^o = \underset{\underline{\mathbf{h}}_{\Phi} \in \mathbb{C}^{N_P N_{CIR}}}{\operatorname{argmin}} \left\{ \sum_{k=0}^{N_{CIR}-1} \sqrt{\sum_{p=0}^{N_P-1} \left|[\underline{\mathbf{h}}_{\Phi}]_{pN_{CIR}+k}\right|^2} \right\},$$

subject to $\|\mathbf{y} - \mathbf{SF}\underline{\mathbf{h}}_{\Phi}\|^2 \leq \sigma^2$.

3) Set $\underline{\mathcal{K}}$ as the support of $\underline{\mathbf{h}}_{\Phi}^o$. The estimate of $\mathcal{K}$ is obtained via $\mathcal{K} = \underline{\mathcal{K}} \cap \mathcal{N}_{CIR}$.

Note that the above algorithm does not recover the values of the entries of $\underline{\mathbf{h}}_{\Phi}^o$, but only its support $\underline{\mathcal{K}}$, from which the support of $\underline{\mathbf{h}}$ is obtained in step 3. Having estimated the sparsity pattern $\mathcal{K}$, we formulate the JMLE under the sparsity

approximation. Let $\underline{\mathbf{h}}_r \in \mathbb{C}^{|\mathcal{K}|}$ denote the vector of DDSF coefficients of *reduced size*, obtained by taking only the entries of $\underline{\mathbf{h}}$ whose indexes are in $\mathcal{K}$. Accordingly, we omit from $\Phi_{(\alpha)}$ the columns whose indexes are not in $\mathcal{K}$, together with the rows whose indexes are not in $\underline{\mathcal{K}}$, and we omit from $\mathbf{F}$ the columns whose indexes are not in $\underline{\mathcal{K}}$. We denote the reduced matrices by $\Phi_{r,(\alpha)} \in \mathbb{C}^{N_P|\mathcal{K}| \times |\mathcal{K}|}$ and $\mathbf{F}_r \in \mathbb{C}^{N_P N_{CIR} \times N_P|\mathcal{K}|}$, respectively. Lastly, denote $\mathbf{H}_{r,(\alpha)} \triangleq \mathbf{SF}_r\Phi_{r,(\alpha)}$, and $\mathbf{G}_{r,(\alpha)} \triangleq \left(\mathbf{H}_{r,(\alpha)}^H \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{H}_{r,(\alpha)}\right)^{-1}$, and approximate the observations via

$$\mathbf{y} \simeq \mathbf{H}_{r,(\alpha)}\underline{\mathbf{h}}_r + \mathbf{w}, \qquad \mathbf{y}^{(aprox.)} \sim \mathcal{CN}\left(\mathbf{H}_{r,(\alpha)}\underline{\mathbf{h}}_r, \mathbf{C}_{\mathbf{w}}\right), \tag{9}$$

which is an approximation as $\underline{\mathbf{h}}$ is assumed only approximately sparse. Using (9), we reformulate the JMLE in (6) as follows:

$$\hat{\alpha}^{CS} = \underset{\alpha \in \mathcal{I}_{1/2}}{\operatorname{argmax}}\left\{\mathbf{y}^H \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{H}_{r,(\alpha)} \mathbf{G}_{r,(\alpha)} \mathbf{H}_{r,(\alpha)}^H \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{y}\right\}, \tag{10a}$$

and

$$\hat{\underline{\mathbf{h}}}_{r,(\hat{\alpha}^{CS})}^{CS} = \mathbf{G}_{r,(\hat{\alpha}^{CS})} \mathbf{H}_{r,(\hat{\alpha}^{CS})}^H \mathbf{C}_{\mathbf{w}}^{-1} \mathbf{y}. \tag{10b}$$

We henceforth refer to the estimator (10) as the *compressed approximated ML estimator* (CS–AML). The performance of the CS–AML may be degraded if $|\mathcal{K}|$ is too large due to the increased number of parameters to be estimated. Thus, we truncate $\mathcal{K}$ based on an energy threshold, by fixing $0 < \epsilon \ll 1$ and selecting $\mathcal{K}$ as the set with minimal cardinality s.t. $\sum_{k \in \mathcal{K}} \sum_{p \in \mathcal{N}_P} |[\underline{\mathbf{h}}_{\Phi}^o]_{pN_{CIR}+k}|^2 \geq (1-\epsilon)\|\underline{\mathbf{h}}_{\Phi}^o\|^2$.

The CS–AML estimates a total of $|\mathcal{K}|+1$ parameters, hence, the number of observations must satisfy $|\mathcal{K}| < N_{obs}$. As $|\mathcal{K}|$ is smaller than N_{CIR}, the CS–AML requires a shorter PS compared to the JMLE and thus, the *spectral efficiency is improved*. Lastly, note that the computational complexity of evaluating (10) can be expressed similarly to that of (6), with the exception that the number of estimated CIR coefficients is decreased from N_{CIR} to $|\mathcal{K}|$. Consequently, the overall complexity of CS-AML is on the order[1] of $\mathcal{O}\left(T \cdot |\mathcal{K}|^3\right)$.

B. Low Complexity Implementation of the CS–AML

Here, we derive a sub-optimal approximation of the CS–AML in which we further reduce the computational complexity by dividing $\mathbf{y}$ into periods, and separately estimating a phase–shifted version of the DDSF coefficients at each period. This estimator is referred to as the *reduced complexity CS–AML* (RCS–AML). Let $\mathbf{S}_p \in \mathbb{C}^{N_0 \times N_{CIR}}$ be a sub-matrix of $\mathbf{S}$ consisting of the entires $[\mathbf{S}]_{n,l}$ with indexes n, l s.t. $pN_0 \leq n < (p+1)N_0$ and $pN_{CIR} \leq l < (p+1)N_{CIR}$. Set $\underline{\mathbf{F}}_r \in \mathbb{C}^{N_{CIR} \times |\mathcal{K}|}$ to be a matrix whose columns are the columns of the matrix $\mathbf{F}_{N_0}^{-1} \otimes \mathbf{I}_{L_{ch}}$ corresponding to the indexes of the set $\mathcal{K}$, and denote $\underline{\mathbf{h}}_{r,p}^{(\alpha)} = \underline{\mathbf{h}}_r e^{j2\pi\alpha p}$, $p \in \mathcal{N}_P$. Lastly, define the vectors $\mathbf{y}_p, \mathbf{w}_p \in \mathbb{C}^{N_0}$ s.t. $[\mathbf{y}_p]_n = [\mathbf{y}]_{pN_0+n}$ and $[\mathbf{w}_p]_n = [\mathbf{w}]_{pN_0+n}$, $p \in \mathcal{N}_P$, $n \in \mathcal{N}_0$. As $w[n]$ is an ACGN, it follows that the covariance matrix of $\mathbf{w}_p$ *does not depend on p*, and we thus denote it by $\underline{\mathbf{C}}_{\mathbf{w}} \triangleq \mathbb{E}\{\mathbf{w}_p\mathbf{w}_p^H\}$. It follows from (9) that $\mathbf{y}_p \simeq \mathbf{S}_p\underline{\mathbf{F}}_r\underline{\mathbf{h}}_{r,p}^{(\alpha)} + \mathbf{w}_p$. Define $\underline{\mathbf{C}}_{r,p} \triangleq (\underline{\mathbf{F}}_r^H \mathbf{S}_p^H \underline{\mathbf{C}}_{\mathbf{w}}^{-1} \mathbf{S}_p \underline{\mathbf{F}}_r)^{-1}$. To facilitate the RCS–AML we assume that $|\mathcal{K}| < N_0$. Under this assumption, for any value of α and for any $p \in \mathcal{N}_P$, *the*

---

[1] Calculating $\underline{\mathbf{h}}_{\Phi}^o$ via the algorithm detailed in [10, Sec. 4.3] requires a complexity on the order of $\mathcal{O}(N_P \cdot N_{CIR})$, hence, we consider the complexity of calculating $\underline{\mathbf{h}}_{\Phi}^o$ negligible compared to $\mathcal{O}\left(T \cdot |\mathcal{K}|^3\right)$.

efficient estimator of $\underline{\mathbf{h}}_{\mathrm{r},p}^{(\alpha)}$ from $\mathbf{y}_p$ is given by[2] [15, Ch. 4.5]:

$$\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)} = \underline{\mathbf{C}}_{\mathrm{r},p}\underline{\mathbf{F}}_{\mathrm{r}}^H\mathbf{S}_p^H\underline{\mathbf{C}}_{\mathbf{w}}^{-1}\mathbf{y}_p, \quad \hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)} \sim \mathcal{CN}\big(\underline{\mathbf{h}}_{\mathrm{r}}e^{j2\pi\alpha p}, \underline{\mathbf{C}}_{\mathrm{r},p}\big). \quad (11)$$

Ignoring the correlation between the estimates $\{\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)}\}_{p\in\mathcal{N}_{\mathrm{p}}}$, we obtain the RCS–AML estimate of $\underline{\mathbf{h}}_{\mathrm{r}}$ and α as the solution to the weighted least squares problem $(\hat{\alpha}^{\mathrm{RCS}}, \hat{\underline{\mathbf{h}}}^{\mathrm{RCS}}) =$

$$\underset{\{\alpha\in\mathcal{I}_{1/2},\underline{\mathbf{h}}_{\mathrm{r}}\in\mathbb{C}^{|\mathcal{K}|}\}}{\mathrm{argmin}} \sum_{p=0}^{N_{\mathrm{p}}-1}(\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)}-e^{j2\pi\alpha p}\underline{\mathbf{h}}_{\mathrm{r}})^H\underline{\mathbf{C}}_{\mathrm{r},p}^{-1}(\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)}-e^{j2\pi\alpha p}\underline{\mathbf{h}}_{\mathrm{r}}).$$

To implement the RCS–AML, define the $|\mathcal{K}|\times|\mathcal{K}|$ matrix $\mathbf{C}_{\mathrm{inv}} \triangleq \big(\sum_{p=0}^{N_{\mathrm{p}}-1}\underline{\mathbf{C}}_{\mathrm{r},p}^{-1}\big)^{-1}$, the function $\mathbf{x}[p]$ s.t. $\mathbf{x}[p] \triangleq \underline{\mathbf{C}}_{\mathrm{r},p}^{-1}\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)}$ for $p\in\mathcal{N}_{\mathrm{p}}$ and $\mathbf{x}[p] = 0$ otherwise, and the discrete-time Fourier transform of $\mathbf{x}[p]$, $\tilde{\mathbf{x}}_{(\alpha)} \triangleq \sum_{p=0}^{N_{\mathrm{p}}-1}\mathbf{x}[p]e^{-j2\pi\alpha p}$. Using these definitions, the RCS–AML can be computed via:

$$\hat{\alpha}^{\mathrm{RCS}} = \underset{\alpha\in\mathcal{I}_{1/2}}{\mathrm{argmax}} \ \tilde{\mathbf{x}}_{(\alpha)}^H\mathbf{C}_{\mathrm{inv}}\tilde{\mathbf{x}}_{(\alpha)}, \quad (12a)$$

$$\hat{\underline{\mathbf{h}}}_{\mathrm{r},(\hat{\alpha}^{\mathrm{RCS}})}^{\mathrm{RCS}} = \mathbf{C}_{\mathrm{inv}}\tilde{\mathbf{x}}_{(\hat{\alpha}^{\mathrm{RCS}})}. \quad (12b)$$

Lastly, we write $\mathbf{C}_{\mathrm{inv}}$ and $\tilde{\mathbf{x}}_{(\alpha)}$ using the scenario parameters: Plugging the definition of $\underline{\mathbf{C}}_{\mathrm{r},p}$, we obtain that $\mathbf{C}_{\mathrm{inv}} = \big(\sum_{p=0}^{N_{\mathrm{p}}-1}\underline{\mathbf{F}}_{\mathrm{r}}^H\mathbf{S}_p^H\underline{\mathbf{C}}_{\mathbf{w}}^{-1}\mathbf{S}_p\underline{\mathbf{F}}_{\mathrm{r}}\big)^{-1}$. Then, using $\underline{\mathbf{C}}_{\mathrm{r},p}$ and (11), $\tilde{\mathbf{x}}_{(\alpha)}$ can be written as $\tilde{\mathbf{x}}_{(\alpha)} = \sum_{p=0}^{N_{\mathrm{p}}-1}\underline{\mathbf{F}}_{\mathrm{r}}^H\mathbf{S}_p^H\underline{\mathbf{C}}_{\mathbf{w}}^{-1}\mathbf{y}_pe^{-j2\pi\alpha p}$.

Note that the RCS–AML is clearly sub-optimal, as the estimates of $\{\hat{\underline{\mathbf{h}}}_{\mathrm{r},p}^{(\alpha)}\}_{p\in\mathcal{N}_{\mathrm{p}}}$ do not utilize the correlation across sequences of length N_0. However, the RCS–AML does not require a matrix inversion at each of the T test points, and its complexity is dominated by the computations of $\tilde{\mathbf{x}}_{(\alpha)}$ and (12a), and is thus of an order of $\mathcal{O}(T\cdot|\mathcal{K}|^2)$, which is considerably lower compared to that of the JMLE (6) and of the CS–AML (10).

IV. NUMERICAL EXAMPLES AND DISCUSSION

In this section we compare the performance of the JMLE (6), the CS–AML (10), and the RCS–AML (12) via a simulations study. We consider an NB-PLC channel, which is a practical periodic channel with an approximately sparse DDSF [13]. The NB-PLC CIR $g[n,l]$ is generated as detailed in [16, Sec. V] with a period of $N_0 = 20$ and a memory of $L_{\mathrm{ch}} = 5$; the CFO is set to $\alpha = 0.2$. The ACGN $w[n]$ is generated according to the noise model detailed in [3] and in the IEEE P1901.2 standard [17], with a set of typical parameters based on [17, Appendix G, LV11]. Sparsity pattern estimation is carried out using the version of the BPDN algorithm detailed in Section III-A with $\sigma^2 = 1.05\cdot\mathbb{E}\{\|\mathbf{w}\|^2\}$ and $\epsilon = 0.005$. For the RCS-AML estimator (12), if $|\mathcal{K}| > N_0$, then $\mathcal{K}$ is truncated to contain only the indexes corresponding to the N_0 most dominant entries of $\underline{\mathbf{h}}$. We consider two values of N_{obs} corresponding to $N_{\mathrm{p}} \in \{4,6\}$. For $N_{\mathrm{p}} = 4$, the number of observations, $N_{\mathrm{obs}} = N_{\mathrm{p}}N_0$, is smaller than the total number of parameters $N_0L_{\mathrm{ch}} + 1$, and thus, the JMLE (6) cannot be

[2]Note that when $|\mathcal{K}| > N_0$, we can always decrease $|\mathcal{K}|$ by taking only the N_0 coefficients with the largest magnitudes out of the $|\mathcal{K}|$ coefficients produced by the sparsity pattern estimation procedure. This truncation clearly results in an increased estimation error for the DDSF coefficients.
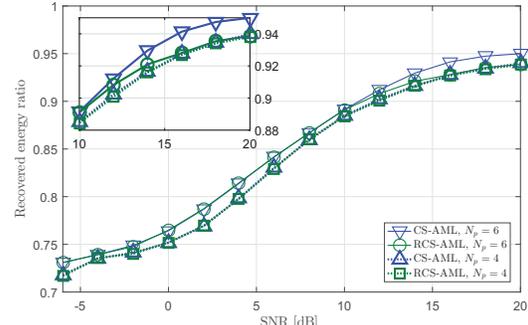

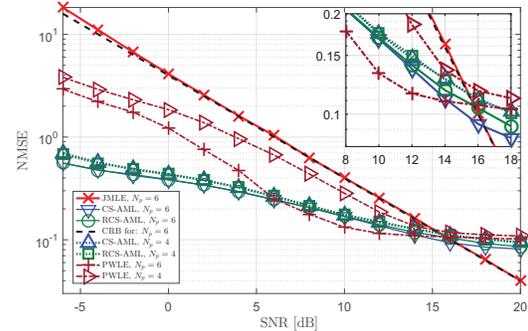Fig. 1.   Performance of sparsity pattern estimation, NB-PLC scenario.

Fig. 2.   CIR NMSE, NB-PLC scenario.

implemented. The estimation of α was implemented using a grid search over $T = 10^4$ equally spaced test points in the interval $\mathcal{I}_{1/2}$. The SNR is defined as $\mathrm{SNR} \triangleq \frac{\|\mathbf{y}-\mathbf{w}\|_2^2}{\mathbb{E}\{\|\mathbf{w}\|_2^2\}}$.

We first evaluate the sparsity pattern estimation in terms of the mean ratio between the CIR energy contained in $\mathcal{K}$ and the total energy of the CIR, evaluated using $r(\mathcal{K}) \triangleq \mathbb{E}_\mathcal{K}\big\{\sum_{k\in\mathcal{K}}|[\underline{\mathbf{h}}]_k|^2\big\}/\|\underline{\mathbf{h}}\|^2$ (note that as $\mathcal{K}$ depends on $\mathbf{y}$, then, $\sum_{k\in\mathcal{K}}|[\underline{\mathbf{h}}]_k|^2$ is a random variable). As observed in Fig. 1, at low SNRs a significant percentage of the estimated DDSF coefficients is set to zero, hence, a significant bias is expected in the estimated CIR. For moderate and high SNRs the indexes of the dominant CIR coefficients are successfully identified for both $N_{\mathrm{p}} = 4$ and for $N_{\mathrm{p}} = 6$. When $N_{\mathrm{p}} = 6$, we observe from Fig. 1 that the mean ratio values of the RCS-AML are slightly smaller compared to the mean ratio achieved by the CS-AML at high SNRs due to the restriction $|\mathcal{K}| \leq N_0$. Note that the term $1 - r(\mathcal{K})$ constitutes a lower bound on the normalized mean-squared error (NMSE) achievable for the estimated CIR. Also note that $r(\mathcal{K})$ is not larger than $\sim 95\%$. The reason is that the DDSF coefficients are only *approximately sparse*, and consequently, the estimator will always set a certain part of the estimated coefficients to zero, which induces an NMSE floor for the estimated CIR coefficients at high SNRs.

Next, we compare the estimation performance for $\underline{\mathbf{h}}$ and α. We also depict the Cramér-Rao bound (CRB) obtained for $N_{\mathrm{p}} = 6$ as a reference (when $N_{\mathrm{p}} = 4$ the CRB does not exist since $N_{\mathrm{p}} < L_{\mathrm{ch}}$ [15, Ch. 7.8]). In addition, we present the estimation performance obtained by adapting the algorithm for joint CFO and CIR estimation proposed in [6] to periodic channels, using a time partitioning approach similarly to [19], referred to as the piecewise LTI estimator (PWLE).

The NMSE for the estimation of $\underline{\mathbf{h}}$, defined as $\mathbb{E}\left\{\|\hat{\underline{\mathbf{h}}} - \underline{\mathbf{h}}\|^2\right\}/\|\underline{\mathbf{h}}\|^2$, is depicted in Fig. 2. As observed in Fig. 2, while the CRB lower bounds the performance of the *unbiased* JMLE, it does not lower bound the performance of the CS-AML and of the RCS-AML, since the CIR estimates obtained using these schemes are biased, due to the fact that the non-zero coefficients excluded from the set $\mathcal{K}$ are estimated as 0. Moreover, since the estimates of the DDSF coefficients are coupled, then decreasing the number of estimated DDSF coefficients results in reducing the MSE of the DDSF coefficients in $\mathcal{K}$. Consequently, the NMSE for the estimated coefficients can be smaller than the CRB, which bounds the NMSE for unbiased estimation of the *entire* DDSF. As observed in Fig. 2, both the CS-AML and the RCS-AML outperform the standard JMLE over a wide range of SNRs, which is achieved thanks to decreasing the number of estimated parameters. Furthermore, even when the number of observed periods is $N_\mathrm{p} = 4$, the CS-AML and the RCS-AML still outperform the JMLE (derived for $N_\mathrm{p} = 6$) at low and medium SNRs. From the above discussion we conclude that *the proposed estimators can provide a substantial performance gain, in addition to increased spectral efficiency and reduced computational complexity*. Additionally, it is observed that for $N_\mathrm{p} = 6$, both the CS-AML and the RCS-AML outperform the PWLE for most of the considered SNR range, for $N_\mathrm{p} = 4$ they are superior for all of the SNR range. This demonstrates the benefits of the CS-AML and the RCS-AML compared to the ad-hoc solution.

The CFO estimation performance is depicted in Fig. 3. As α is periodic, its estimation performance is measured via the mean-squared periodic error (MSPE), defined as $\mathrm{MSPE}(\hat{\alpha} - \alpha) \triangleq \mathbb{E}\left\{\langle\hat{\alpha} - \alpha\rangle^2\right\}$ [18, Sec. II]. Note that at high SNRs the CRB also lower bounds the MSPE [18]. It is observed in Fig. 3 that for $N_\mathrm{p} = 6$, both the CS-AML and the RCS-AML outperform the JMLE over a wide range of SNRs. This can be explained by noting that the CFO estimation in (6a), (10a), and (12a), yields the values of α maximizing the energy of a projection of the observations $\mathbf{y}$ onto a signal subspace which depends on the tested value of α. The CS-AML and the RCS-AML only utilize signal space dimensions which account for dominant DDSF coefficients, resulting in improved performance. Furthermore, as the estimates of α and $\underline{\mathbf{h}}$ are coupled, [7, Sec. 2.], the bias in the estimation of $\underline{\mathbf{h}}$ can result in estimates of α with an MSPE lower than the CRB. Lastly we observe that both the CS-AML and the RCS-AML obtain greatly improved CFO estimation performance compared to the PWLE over a wide range of SNRs.

V. Conclusions

In this paper we proposed an estimation scheme for the joint estimation of the CIR and CFO over channels with periodic characteristics. The proposed estimators exploit the approximately sparse nature of the CIR using compression–based techniques, and can be implemented with a reduced computational complexity at the cost of a minor performance degradation. The performance of the proposed schemes is evaluated in a simulation study corresponding to a practical scenario, and compared to the performance of the JMLE and of a reference ad-hoc estimation scheme based on adapting
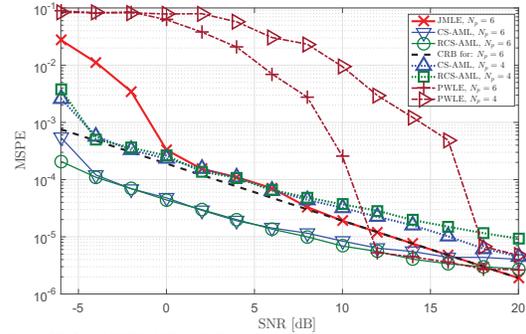

Fig. 3. CFO MSPE, NB-PLC scenario.

an estimation scheme for LTI channels. The numerical results illustrate that properly exploiting the periodicity and the approximate sparsity of the periodic channel is a key for obtaining improved estimation performance.

References

[1] W. A. Gardner, A. Napolitano, and L. Paura. "Cyclostationarity: Half a century of research". *Signal Processing*, vol. 86, Apr. 2006, pp. 639-697.

[2] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. "What will 5G be?". *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, Jun. 2014, pp. 1065-1082.

[3] M. Nassar, J. Lin, Y. Mortazavi, A. Dabak, I. H. Kim, and B. L. Evans. "Local utility power line communications in the 3–500 kHz band: Channel impairments, noise, and standards". *IEEE Signal Process. Mag.*, vol. 29, no. 5, Aug. 2012, pp. 116-127.

[4] W. U. Bajwa, A. M. Sayeed and R. Nowak. "Learning sparse doubly-selective channels". *Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL, 2008, pp. 575-582.

[5] Y. Jiang, X. Zhu, E. G. Lim, Y. Huang, and H. Lin. *Semi-Blind Carrier Frequency Offset Estimation and Channel Equalization*. Springer, 2015.

[6] M. Morelli and U. Mengali. "Carrier-frequency estimation for transmissions over selective channels". *IEEE Trans. Commun.*, vol. 48, no. 9, Sep. 2000, pp. 1580-1589.

[7] M. Ghogho and A. Swami. "Training design for multipath channel and frequency-offset estimation in MIMO systems". *IEEE Trans. Signal Process.*, vol. 54, no. 10, Oct. 2006, pp. 3957-3965.

[8] R. Shaked, N. Shlezinger, and R. Dabora. "Carrier frequency offset estimation for linear channels with periodic characteristics". *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Edinburgh, UK, 2016, pp. 1-5.

[9] A. M. Bruckstein, M. Alfred, D. L. Donoho, and M. Elad. "From sparse solutions of systems of equations to sparse modeling of signals and images". *SIAM Review*, vol. 51, no. 1, Feb. 2009, pp. 34-81.

[10] E. V. D. Berg, M. Schmidt, M. P. Freidlender, and K. Murphy. "Group sparsity via linear-time projection". Technical Report, Department of Computer Science, University of British Columbia, 2008.

[11] Y. Xu, J. Zhong, Y. Cai, M. Zhao and B. Chen. "Joint CFO and sparse channel estimation for MIMO-OFDM systems via the SAGE algorithm". *International Conference on Wireless Communications and Signal Processing (WCSP)*, Hangzhou, China, Oct. 2013, pp. 1-5.

[12] R. F. Senyuva, G. K. Kurt, and E. Anarim. "Compressive sensing based joint frequency offset and channel estimation for OFDM". *EURASIP J. Wireless Com. Network.*, Mar. 2016, pp. 1-13.

[13] F. Gianaroli, F. Pancaldi, and G. M. Vitetta. "On the use of Zadeh's series expansion for modeling and estimation of indoor powerline channels". *IEEE Trans. Commun.*, vol. 62, no. 7, Jul. 2014, pp. 2558-2568.

[14] A. V. Oppenheim, R. W. Schafer, and J. R. Buck. *Discrete-Time Signal Processing*. Prentice-Hall, 1998.

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

[16] N. Shlezinger and R. Dabora. "On the capacity of narrowband PLC channels". *IEEE Trans. Commun.*, vol. 63, no. 4, Apr. 2015, pp. 1191-1201.

[17] IEEE Standards Association. "P1901.2/D0.09.00 draft standard for low frequency (less than 500 kHz) narrow band power line communications for smart grid applications". Jun. 2013.

[18] T. Routtenberg and J. Tabrikian. "Non-Bayesian periodic Cramér-Rao bound". *IEEE Trans. Signal Process.*, vol. 61, no. 4, Feb. 2013, pp. 1019-1032.

[19] M. A. Tunc, E. Perrins, and L. Lampe. "Optimal LPTV-aware bit loading in broadband PLC". *IEEE Trans. Commun.*, vol. 61, no. 12, Dec. 2013, pp. 5152-5162.