Computers and Electrical Engineering 72 (2018) 406–419

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

On channel estimation for power line communication systems in the presence of impulsive noise☆

Deep Shrestha[a,*], Xavier Mestre[b], Miquel Payaró[b]

[a] Signal Theory and Communications Department, Universitat Politècnica de Catalunya (UPC), C/Jordi Girona 1–3, Building D4, Barcelona, 08034, Spain

[b] Centre Tecnològic de Telecomunicacions de Catalunya (CTTC/CERCA), Parc Mediterrani de la Tecnologia (PMT), Building B4, Av. Carl Friedrich Gauss 7, Castelldefels, 08860, Spain

ARTICLE INFO

Article history:
Received 29 December 2017
Revised 3 October 2018
Accepted 4 October 2018
Available online 12 October 2018

Keywords:
Channel estimation
Impulsive noise
Power line communication
Maximum-likelihood estimation

ABSTRACT

In this paper, two maximum-likelihood (ML) channel estimators that are robust and perform consistent channel estimation in the presence of impulsive noise (IN) for power line communication (PLC) systems are proposed and analyzed. The two estimators differ on the way they exploit the estimated IN to determine the coefficients of the channel impulse response. In the first approach, the channel estimator treats the estimated IN as a deterministic quantity; in the second as a random quantity. The performances of both estimators are analyzed and numerically evaluated. As it will be shown, between the two proposed estimators, the one that is based on the random approach outperforms the deterministic one in all typical PLC scenarios. However, the deterministic approach-based estimator can perform consistent channel estimation regardless of the IN behavior with less computational effort and becomes an efficient channel estimation strategy in situations where high computational complexity cannot be afforded.

1. Introduction

The smart grid envisions the electrical power distribution network as a physical infrastructure that relies on communication technologies to efficiently deliver the generated power to the end consumers [1–3]. Due to the extensive outreach of power lines in electric grids, using them also for the purpose of data transmission is considered as an economically efficient way to support the communication requirements of smart grid applications [4]. In order to serve as a communication backbone for the smart grid, power line communication (PLC) however has to overcome many impairments that make the power line environment harsh for reliable data transmission [5,6].

One of the major drawbacks for PLC is the power line itself. Power lines were originally designed to transmit electrical energy only and severely impair communication signals bearing frequencies higher than the fundamental operating frequency of the electrical signal (50 Hz or 60 Hz). In a typical power line network, where there are multiple branches and impedance mismatches along an electrical path, a communication signal transmitted through it is typically reflected multiple times before reaching the receiving node [5]. As a result, the transmitted communication signal is received via multiple paths and is severely affected by the frequency selectivity of the channel [7]. In order to overcome the resulting frequency

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Area Editor Dr. E. Cabal-Yepez.

* Corresponding author.
E-mail addresses: deep.shrestha@upc.edu, er.dipshrestha@gmail.com (D. Shrestha), xmestre@cttc.es (X. Mestre), mpayaro@cttc.es (M. Payaró).

https://doi.org/10.1016/j.compeleceng.2018.10.006
0045-7906/

selectivity, modern PLC systems exploit multicarrier modulation techniques based on orthogonal frequency division multiplexing (OFDM) [8]. The simple single-tap equalization feature of OFDM allows easy equalization of the symbols against the frequency selectivity of the power line channel, provided the channel is accurately estimated [9]. Typically in order to estimate the channel, pilot symbols that are also known to the receiver are transmitted. Upon receiving the signal, the receiver exploits the known pilot symbols to estimate the channel coefficients and use them to equalize the received symbols.

Apart from the channel frequency selectivity, randomly occurring impulsive noise (IN) in the power line environment also hinders reliable data transmission. Due to the operation of the power converters along with other essential electrical components that are used for the efficient delivery of electrical power, random IN inevitably arises in the electrical network [1]. The brief duration of the IN in time domain leads to a wide band effect in the frequency domain. As a result, when such noise occurs during data transmission, the energy of IN spreads among all the subcarriers of an OFDM based PLC system. The spread of energy takes place while demodulating the received signal by performing fast Fourier transformation (FFT) on the time domain samples that are affected by the IN [10]. In such a hostile environment, the pilot symbols used to estimate the channel are also severely affected, making precise channel estimation a challenging task [11].

The problem of estimating the channel in the presence of random IN for an OFDM based PLC system has gained a lot of attention in recent years [9,11–16]. Conventional approaches follow least-squares (LS) channel frequency response (CFR) estimation, typically in conjunction with the IN mitigation scheme. The LS estimator which shows optimum performance in the presence of additive white Gaussian noise (AWGN) tends to loose its optimality in the presence of IN. Therefore, in order to exploit LS estimation of CFR coefficients, channel estimation schemes in [11–14] exploit a threshold dependent nulling technique to reduce the power of the IN samples present in the received signal before estimating the channel. However, due to the non-linearity of the IN mitigation scheme, some samples of the IN might still be present in the received signal. In such a situation, the estimate of the CFR obtained as an output of the LS channel estimator is not accurate [17]. To enhance the performance of channel estimation schemes based on LS in conjunction with non-linear IN mitigation scheme, denoising algorithms that reduce noise from the estimated coefficients of the CFR are proposed in [9,15,16]. In order to remove noise from the estimated CFR coefficients, the proposed algorithms typically follow domain transformation techniques. As a first step, the estimated coefficients of the channel in frequency domain are transformed into time domain to generate the impulse response of the channel. After the transformation, the significant taps that bear most of the channel energy are then determined based on a threshold value. The remaining samples in the channel impulse response (CIR) are nulled and a new CFR is generated based on the denoised impulse response of the channel. However, the use of such threshold dependent approach to identify the most significant channel taps in time domain shows inconsistent behavior in PLC environments with IN. Hence, channel estimation strategies that are not only robust, but also consistent in harsh power line environment are required to exploit PLC based technologies to support smart grid applications.

In order to precisely estimate the channel in the harsh power line environment where the IN is prominent, a better performance than the conventional channel estimation strategies can be achieved, if instead of trying to suppress the IN affected samples in the received signal, the IN is estimated and used in the process of the estimation of the channel [18]. To do so, the sparse IN present in the received signal can be estimated by using the signal processing schemes like the ones proposed in [19–21] can be used. Exploiting the observations, in terms of estimated IN and the received signal, a likelihood function can be defined. The channel coefficients can now be determined by estimating them as parameter values that maximize the likelihood function.

Maximum-likelihood (ML) estimation of the channel by exploiting the estimated IN for PLC systems has not yet been widely explored. In one of our prior works, reported in [18], we showed some preliminary results on how the estimated IN samples can be used to estimate the channel for PLC. The results show that channel can be estimated with better precision in the presence of IN by exploiting the ML technique. To this end, in this article we propose two ML channel estimators that exploit the estimated IN to robustly estimate the CIR in an impulsive-noise power line environment for OFDM based PLC systems and study their performance. In order to formulate the ML channel estimators, we begin by assuming that the IN is estimated and is available to the receiver for further exploitation. The estimated IN, in terms of power and the IN sample locations, is then further exploited by taking two different approaches to derive two ML channel estimators. In the first approach, the estimated IN samples are treated as a deterministic quantity and an ML CIR estimator in closed form is derived. In the second approach, the estimated IN is considered as a random quantity and an ML channel estimator that is based on an iterative approximation of the CIR coefficients using a robust likelihood function is derived. Furthermore, the performances of the proposed estimators in typical PLC scenarios are analyzed and the variances of the estimation errors along with their respective Cramer-Rao bounds are evaluated. In order to assess the performance of the proposed estimators in scenarios where errors are made during IN support estimation, the performances of both estimators are also analyzed when the IN support is underestimated and when the support is overestimated. As it will be shown, based on the numerical validation of the performance of the proposed estimators, the random ML estimator outperforms the deterministic ML estimator in all typical PLC scenarios. However, the deterministic ML estimator performs consistent estimation of the CIR coefficients with significantly less computational effort than the random ML estimator, at the expense of robustness.

This document is divided into seven sections. In Section 2, we briefly outline the OFDM based PLC system model that is considered in this work. In Section 3, both ML channel estimators are derived and the variances of both estimators are evaluated in Section 4. The determination of Cramer-Rao bounds (CRBs) of both estimators along with numerical validation of the performance of proposed estimators are done in Section 5, for which the simulation results are presented in Section 6. Finally, the paper concludes with a brief summary in Section 7.

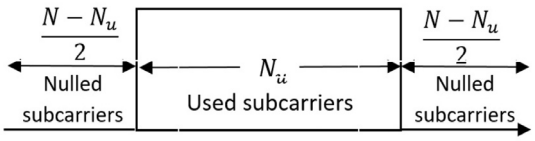

Fig. 1. Locations of used and unused subcarriers in the spectrum.

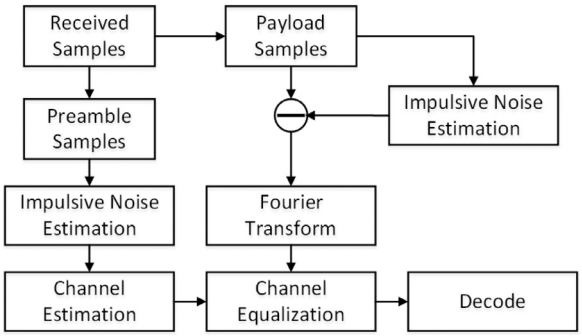

Fig. 2. Block diagram of proposed channel estimation scheme.

2. System model

In this work, we consider an OFDM based PLC system with N subcarriers. Among the available N subcarriers, only N_u subcarriers are used for the data transmission. The rest $N - N_u$ subcarriers are nulled/unused as shown in Fig. 1. A transmission frame consisting of one preamble symbol (containing the known pilot symbols) followed by a payload symbol is considered. A static channel for the duration of a frame transmission is also assumed such that the channel estimated exploiting the known symbols in the preamble is used by the receiver to the equalize symbols received in the payload as shown in Fig. 2.

Taking into account the effect of IN occurring during signal transmission, the time domain received signal corresponding to the preamble is expressed as:

$$\mathbf{y} = \sqrt{N}\mathbf{F}\mathbf{E}\mathbf{\Lambda}_p\mathbf{G}^H\mathbf{h} + \mathbf{i}_p + \mathbf{w}_p. \tag{1}$$

The variable $\mathbf{y}$ in Eq. (1) is a vector of dimension $N \times 1$ and contains the time domain samples of the complex baseband equivalent received signal. The variable $\mathbf{F}$ in Eq. (1) denotes the unitary inverse-discrete Fourier transform (IDFT) matrix of dimension $N \times N$, whose elements are defined as $[\mathbf{F}]_{a,b} = \frac{1}{\sqrt{N}} \exp^{j2\pi ab/N}$, where a and b denote the row and column indexes of the matrix $\mathbf{F}$. The variable $\mathbf{E}$ in Eq. (1) denotes a selection matrix of dimension $N \times N_u$ that has ones at the entries identifying the locations of the subcarriers that are used for the data transmission. $\mathbf{\Lambda}_p$ in Eq. (1) denotes a diagonal matrix of dimension $N_u \times N_u$ whose diagonal entries contain the frequency domain pilot symbols transmitted in the preamble. The matrix $\mathbf{G}$ in Eq. (1) is a section of Fourier matrix $\mathbf{F}$ that is obtained by selecting the N_u columns of $\mathbf{F}$ that are associated with the pilot positions and the corresponding L upper rows. The channel is denoted by the variable $\mathbf{h}$ in Eq. (1), which is a vector of dimension $L \times 1$, L being the length of the CIR. The variables $\mathbf{i}_p$ and $\mathbf{w}_p$ in Eq. (1), each of dimension $N \times 1$, denote vectors bearing the time domain samples of IN and background noise, respectively.

The samples of the background noise are assumed to be i.i.d. additive white Gaussian noise (AWGN) random variables. The entries of vector $\mathbf{w}_p$ are therefore Gaussian distributed random variables having zero mean and variance $\bar{\sigma}_w^2$. The IN vector $\mathbf{i}_p$ is a sparse vector, having only N_{imp} non-zero entries. Furthermore, the indexes within the set $\{1, \ldots, N\}$ that are contaminated by IN are denoted by $\mathcal{I} = \left\{n_1, \ldots, n_{N_{\text{imp}}}\right\}$ such that $N_{\text{imp}} = |\mathcal{I}|$. The signal to noise ratio (SNR) of the system is defined as the ratio $\bar{\sigma}_s^2/\bar{\sigma}_w^2$, where $\bar{\sigma}_s^2$ is the power of the transmitted signal and the IN to background noise power ratio (INR) is defined as the ratio $\bar{\sigma}_i^2/\bar{\sigma}_w^2$, where $\bar{\sigma}_i^2$ is the power of the IN samples.

3. Maximum-likelihood channel estimation

In this section, we present the derivations of the proposed ML channel estimators. After receiving the samples transmitted during the preamble transmission time, as shown in Eq. (1), the estimated IN and the known pilot symbols are exploited to determine the ML estimate of the CIR coefficients. While exploiting the estimated IN, as mentioned before in Section 1, we consider two different approaches and outline them in the following two subsections. In Section 3.1, we derive the ML channel estimator that is based on the deterministic approach. In Section 3.2, the ML channel estimator that follows the random approach is derived.

3.1. Deterministic approach

Under the deterministic setting, we express the time domain received signal corresponding to preamble in Eq. (1) as

$$\mathbf{y} = \sqrt{N}\mathbf{F}\mathbf{E}\mathbf{\Lambda}_p\mathbf{G}^H\mathbf{h} + \mathbf{U}\mathbf{i}_c + \mathbf{w}_p, \tag{2}$$

where $\mathbf{i}_c$ is a non-sparse vector containing only the non-zero entries of $\mathbf{i}_p$ and

$$\mathbf{U} = \begin{bmatrix} \mathbf{e}_{n_1} & \cdots & \mathbf{e}_{n_{N_{\text{imp}}}} \end{bmatrix}.$$

The matrix $\mathbf{U}$ in Eq. (2) is of dimension $N \times N_{\text{imp}}$, where each column vector $\mathbf{e}_{\mathcal{I}}$ is of length $N \times 1$ and has a single 1 identifying the location of IN sample as indexed by $\mathcal{I}$. Furthermore, we also define the matrix $\mathbf{P}$ as a diagonal matrix whose entries are ones only at the locations where there are IN samples. In other words, $\mathbf{P} = \mathbf{U}\mathbf{U}^T$.

Assuming deterministic IN samples, the negative log-likelihood function of Eq. (2) takes the form

$$\zeta_{DML} = N \log \sigma_w^2 + \frac{1}{\sigma_w^2} \left\| \mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h} - \mathbf{U}\mathbf{i}_c \right\|^2, \tag{3}$$

where $\Theta = \mathbf{F}\mathbf{E}\mathbf{\Lambda}_p\mathbf{G}^H$, which will be used through out the remainder of the paper. Provided the support of the IN is known, the likelihood function in Eq. (3) can be equivalently reformulated, only as a function of $\mathbf{h}$, as:

$$\zeta_{DML}(\mathbf{h}) = \frac{1}{N - N_{\text{imp}}} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h} \right)^H \mathbf{P}^{\perp} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h} \right), \tag{4}$$

where $\mathbf{P}^{\perp} = \mathbf{I}_N - \mathbf{P} = \mathbf{I}_N - \sum_{i=1}^{N_{\text{imp}}} \mathbf{e}_{n_i}\mathbf{e}_{n_i}^T$, $\mathbf{I}_N$ being the identity matrix of dimension $N \times N$, is a diagonal matrix containing ones in its diagonal entries corresponding to the locations of the samples in the received signal that are not impaired by the IN.

Taking the derivative of Eq. (4) with respect to $\mathbf{h}$ and equating it to zero, assuming that the supports of the IN samples are provided by a previous inference procedure (e.g., the one reported in [21]), the deterministic ML channel estimator can be expressed as:

$$\hat{\mathbf{h}}_{DML} = \frac{1}{\sqrt{N}} \left(\Theta^H \hat{\mathbf{P}}^{\perp} \Theta \right)^{-1} \Theta^H \hat{\mathbf{P}}^{\perp} \mathbf{y}, \tag{5}$$

where $\hat{\mathbf{P}}^{\perp} = \mathbf{I}_N - \hat{\mathbf{P}}$ is the estimated value of $\mathbf{P}^{\perp}$ and $\hat{\mathbf{P}}$ is the estimated value of $\mathbf{P}$.

3.2. Random approach

Under the random setting, the time domain received signal model in Eq. (1) is expressed as:

$$\mathbf{y} = \sqrt{N}\Theta\mathbf{h} + \mathbf{n}, \tag{6}$$

where $\mathbf{n} = \mathbf{i}_p + \mathbf{w}_p$. Furthermore, we assume that the noise follows a zero mean circular Gaussian distribution $\mathbf{n} \sim \mathcal{CN}(0, \mathbf{C}_n)$, where the covariance matrix has the structure

$$\mathbf{C}_n = \sigma^2\mathbf{P} + \sigma_w^2\mathbf{P}^{\perp},$$

where $\sigma^2 = \sigma_i^2 + \sigma_w^2$ denotes the cumulative noise power at the locations of the IN samples. The negative log-likelihood function of Eq. (6), assuming random IN, takes the form

$$\zeta_{RML} = \log |\mathbf{C}_n| + \bar{\mathbf{y}}^H \mathbf{C}_n^{-1} \bar{\mathbf{y}}, \tag{7}$$

where $\bar{\mathbf{y}} = \mathbf{y} - \sqrt{N}\Theta\mathbf{h}$.

Assuming that the support of the IN is previously estimated, the resulting noise covariance matrix can be constructed as $\hat{\mathbf{C}}_n = \hat{\sigma}^2\hat{\mathbf{P}} + \hat{\sigma}_w^2\hat{\mathbf{P}}^{\perp}$, where $\text{tr}[\hat{\mathbf{P}}] = \hat{N}_{\text{imp}}$ and $\text{tr}[\hat{\mathbf{P}}^{\perp}] = N - \hat{N}_{\text{imp}}$. Differentiating Eq. (7) with respect to the cumulative noise power and the background noise power and equating them to zero, the cumulative noise power and the background noise power in terms of locations of IN samples and its orthogonal locations and the received signal can be expressed as:

$$\hat{\sigma}^2 = \frac{1}{\hat{N}_{\text{imp}}} \bar{\mathbf{y}}^H \hat{\mathbf{P}} \bar{\mathbf{y}}$$

and

$$\hat{\sigma}_w^2 = \frac{1}{N - \hat{N}_{\text{imp}}} \bar{\mathbf{y}}^H \hat{\mathbf{P}}^{\perp} \bar{\mathbf{y}}. \tag{8}$$

Furthermore, plugging the values of estimated cumulative noise power and the background noise power into Eq. (7), the normalized likelihood function for the random approach, only in terms of $\mathbf{h}$, can be expressed as:

$$\zeta_{RML}(\mathbf{h}) = \frac{\hat{N}_{\text{imp}}}{N} \log \left(\frac{1}{\hat{N}_{\text{imp}}} \bar{\mathbf{y}}^H \hat{\mathbf{P}} \bar{\mathbf{y}} \right) + \frac{N - \hat{N}_{\text{imp}}}{N} \log \left(\frac{1}{N - \hat{N}_{\text{imp}}} \bar{\mathbf{y}}^H \hat{\mathbf{P}}^{\perp} \bar{\mathbf{y}} \right). \tag{9}$$

The ML estimate of the CIR is now achieved by finding the optimum of the likelihood function in Eq. (9). In order to do so, we solve $\nabla\zeta_{RML}(\mathbf{h}) = \mathbf{0}$ by applying the Newton-Raphson method [22]. Initiating the solution finding procedure with an initial guess $\mathbf{h}_0$ and exploiting the Hessian and the gradient of the likelihood function, an iterative approximation of the root can be computed as:

$$\mathbf{h}_{k+1} = \mathbf{h}_k - [\mathbb{H}\zeta_{RML}(\mathbf{h}_k)]^{-1}\nabla\zeta_{RML}(\mathbf{h}_k), \tag{10}$$

where $\mathbb{H}\zeta_{RML}$ is the Hessian matrix and $\nabla\zeta_{RML}$ is the gradient of the likelihood function. The final estimate of the CIR is achieved when $\frac{||\mathbf{h}_{k+1}-\mathbf{h}_k||^2}{||\mathbf{h}_k||^2} \leq t$, where t is a threshold that defines the stopping criterion. (*The reader is referred to Appendix for the consistency check of the likelihood function in* Eq. (9) as $N \to \infty$.)

In order to proceed with the iterative-approximation of the channel estimate, we begin by evaluating the Hessian matrix and the gradient of the cost function with respect to $\mathbf{h}$.

3.2.1. Calculation of Hessian and gradient of cost function

To evaluate the gradient of the likelihood function in Eq. (9), we begin by evaluating the complex derivatives that establish

$$\nabla\zeta_{RML}(\mathbf{h}) = \frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathcal{R}(\mathbf{h})} = \widetilde{\mathbf{\Xi}}\begin{bmatrix}\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*}\\ \frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}}\end{bmatrix}, \tag{11}$$

where $\mathcal{R}(\mathbf{h}) = \left[\mathrm{Re}^T(\mathbf{h}), \mathrm{Im}^T(\mathbf{h})\right]^T$, $\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}}$ denotes derivative of the cost function with respect to the CIR, $\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*}$ is the derivative of the cost function with respect to the complex conjugate of the CIR and $\widetilde{\mathbf{\Xi}} = \begin{bmatrix}\mathbf{I}_L & \mathbf{I}_L\\ -\mathbf{j}\mathbf{I}_L & \mathbf{j}\mathbf{I}_L\end{bmatrix}$. Furthermore, the Hessian of the cost function can be calculated by evaluating

$$\mathbb{H}\zeta_{RML}(\mathbf{h}) = \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathcal{R}(\mathbf{h})\partial\mathcal{R}^T(\mathbf{h})} = \widetilde{\mathbf{\Xi}}\begin{bmatrix}\frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*\partial\mathbf{h}^H} & \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*\partial\mathbf{h}^T}\\ \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}\partial\mathbf{h}^H} & \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}\partial\mathbf{h}^T}\end{bmatrix}\widetilde{\mathbf{\Xi}}^T. \tag{12}$$

To simplify the notation, we express the Hessian matrix equivalently as

$$\mathbb{H}\zeta_{RML}(\mathbf{h}) = \widetilde{\mathbf{\Xi}}\begin{bmatrix}\bar{\mathbb{H}}_c(\mathbf{h}) & \mathbb{H}_c(\mathbf{h})\\ (\mathbb{H}_c(\mathbf{h}))^T & \left(\bar{\mathbb{H}}_c(\mathbf{h})\right)^*\end{bmatrix}\widetilde{\mathbf{\Xi}}^T, \tag{13}$$

where $\mathbb{H}_c(\mathbf{h}) = \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*\partial\mathbf{h}^T}$ and $\bar{\mathbb{H}}_c(\mathbf{h}) = \frac{\partial^2\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*\partial\mathbf{h}^H}$.

Proceeding with the calculations of the complex derivatives, the elements of the gradient vector and the Hessian matrix can be evaluated as:

$$\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*} = -\frac{\hat{N}_{\mathrm{imp}}}{\sqrt{N}}\frac{\mathbf{\Theta}^H\hat{\mathbf{P}}\bar{\mathbf{y}}}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}\bar{\mathbf{y}}} - \frac{N-\hat{N}_{\mathrm{imp}}}{\sqrt{N}}\frac{\mathbf{\Theta}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}}, \tag{14}$$

$$\mathbb{H}_c(\mathbf{h}) = \frac{\hat{N}_{\mathrm{imp}}}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}\bar{\mathbf{y}}}\mathbf{\Theta}^H\left(\hat{\mathbf{P}} - \frac{\hat{\mathbf{P}}\bar{\mathbf{y}}\bar{\mathbf{y}}^H\hat{\mathbf{P}}}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}\bar{\mathbf{y}}}\right)\mathbf{\Theta} + \frac{N-\hat{N}_{\mathrm{imp}}}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}}\mathbf{\Theta}^H\left(\hat{\mathbf{P}}^\perp - \frac{\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp}{\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}}\right)\mathbf{\Theta}, \tag{15}$$

and

$$\bar{\mathbb{H}}_c(\mathbf{h}) = -\hat{N}_{\mathrm{imp}}\frac{\mathbf{\Theta}^H\hat{\mathbf{P}}\bar{\mathbf{y}}\bar{\mathbf{y}}^H\hat{\mathbf{P}}\mathbf{\Theta}^*}{\left(\bar{\mathbf{y}}^H\hat{\mathbf{P}}\bar{\mathbf{y}}\right)^2} - \left(N-\hat{N}_{\mathrm{imp}}\right)\frac{\mathbf{\Theta}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\mathbf{\Theta}^*}{\left(\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}\right)^2}, \tag{16}$$

where we recall that $\bar{\mathbf{y}} = \mathbf{y} - \sqrt{N}\Theta\mathbf{h}$. Upon evaluating Eqs. (14)–(16), and plugging them into Eqs. (11) and (13), the gradient and the Hessian of the cost can be readily obtained.

4. Asymptotic performance characterization

In this section we evaluate the error in the channel estimation of the proposed methods. We divide this section into two subsections, one for each approach, and outline in detail the characterizations of the variances of the estimation errors of both estimators. In the first subsection, we characterize the channel estimation error of the deterministic ML estimator. In the second subsection, the variance of the channel estimation error of the random ML estimator is evaluated taking an asymptotic approach, when $N \to \infty$.

4.1. Variance of the deterministic ML estimator

To calculate the variance of the deterministic ML channel estimator, we recall Eq. (5) and replace $\mathbf{y} = \sqrt{N}\mathbf{F}\mathbf{E}\mathbf{\Lambda}_p\mathbf{G}^H\bar{\mathbf{h}} + \bar{\mathbf{n}}$, where $\bar{\mathbf{h}}$ denotes the true channel and $\bar{\mathbf{n}}$ denotes the true noise samples, we observe that

$$\hat{\mathbf{h}}_{DML} = \bar{\mathbf{h}} + \frac{1}{\sqrt{N}}\left(\mathbf{\Theta}^H\hat{\mathbf{P}}^{\perp}\mathbf{\Theta}\right)^{-1}\mathbf{\Theta}^H\hat{\mathbf{P}}^{\perp}\bar{\mathbf{n}}. \tag{17}$$

The true noise in Eq. (17) is assumed to follow $\bar{\mathbf{n}} \sim \mathcal{CN}\left(\mathbf{0}, \bar{\mathbf{C}}_n\right)$ distribution. Similar to $\mathbf{C}_n$, the true noise covariance, $\bar{\mathbf{C}}_n$, is defined as $\bar{\mathbf{C}}_n = \bar{\sigma}^2\bar{\mathbf{P}} + \bar{\sigma}_w^2\bar{\mathbf{P}}^{\perp}$, where $\bar{\sigma}^2$ is the true variance of the IN, $\bar{\mathbf{P}}$ is the diagonal matrix with ones as its entries at the locations identifying the true support of the IN samples, $\bar{\sigma}_w^2$ is the true variance of the background noise and $\bar{\mathbf{P}}^{\perp} = \mathbf{I}_N - \bar{\mathbf{P}}$ is the diagonal matrix with ones at its entries identifying the samples in the received signal that are not corrupted by the IN. Note that $\hat{\mathbf{P}} = \bar{\mathbf{P}}$ and $\hat{\mathbf{P}}^{\perp} = \bar{\mathbf{P}}^{\perp}$, only if the estimated support of the IN aligns with the true one.

Trivially from Eq. (17), we see that $\hat{\mathbf{h}}_{DML}$ is unbiased regardless of the IN behavior and $\sqrt{N}\left(\hat{\mathbf{h}}_{DML} - \bar{\mathbf{h}}\right)$ is Gaussian distributed with zero mean and covariance

$$\mathbf{C}_{DML} = \left(\mathbf{\Theta}^H\hat{\mathbf{P}}^{\perp}\mathbf{\Theta}\right)^{-1}\mathbf{\Theta}^H\hat{\mathbf{P}}^{\perp}\bar{\mathbf{C}}_n\hat{\mathbf{P}}^{\perp}\mathbf{\Theta}\left(\mathbf{\Theta}^H\hat{\mathbf{P}}^{\perp}\mathbf{\Theta}\right)^{-1}. \tag{18}$$

4.2. Variance of the random ML estimator

To evaluate the variance of the random ML estimator, we take an asymptotic approach and characterize the behavior of the estimation error by assuming large sample size N. In order to do so, we recall the cost function $\zeta_{RML}(\mathbf{h})$ derived in Eq. (9) and evaluate its gradient around the true channel $\bar{\mathbf{h}}$. The Taylor series expansion of the gradient of the cost function can be expressed as

$$\nabla\zeta_{RML}(\mathbf{h}) = \left[\nabla\zeta_{RML}(\mathbf{h})\right]_{\mathbf{h}=\bar{\mathbf{h}}} + \left[\mathbb{H}\zeta_{RML}(\mathbf{h})\right]_{\mathbf{h}=\bar{\mathbf{h}}}\mathcal{R}\left(\mathbf{h} - \bar{\mathbf{h}}\right) + o_p(1), \tag{19}$$

upon neglecting the higher order terms and evaluating equation Eq. (19) at $\mathbf{h} = \hat{\mathbf{h}}_{RML}$, we see that

$$\mathcal{R}\left(\hat{\mathbf{h}}_{RML} - \bar{\mathbf{h}}\right) = -\left[\mathbb{H}_{\zeta_{RML}}(\mathbf{h})\right]^{-1}_{\mathbf{h}=\bar{\mathbf{h}}}\left[\nabla\zeta_{RML}(\mathbf{h})\right]_{\mathbf{h}=\bar{\mathbf{h}}}, \tag{20}$$

or equivalently,

$$\hat{\mathbf{h}}_{RML} - \bar{\mathbf{h}} = -\mathbf{\Xi}\left[\mathbb{H}\zeta_{RML}(\mathbf{h})\right]^{-1}_{\mathbf{h}=\bar{\mathbf{h}}}\left[\nabla\zeta_{RML}(\mathbf{h})\right]_{\mathbf{h}=\bar{\mathbf{h}}}, \tag{21}$$

where we have used $\mathbf{\Xi} = [\mathbf{I}_L, \mathrm{j}\mathbf{I}_N]$.

In order to establish the asymptotic behavior of the large sample size error, we first define the convergence of the Hessian matrix and then proceed with the characterization of the asymptotic distribution of the cost function gradient.

4.2.1. Convergence of the Hessian matrix

To determine the convergence of the Hessian matrix derived in Eq. (13), we evaluate its eigenvalues by studying the bilinear form

$$\mathcal{R}^T(\mathbf{a})\mathbb{H}\zeta_{RML}(\mathbf{h})\mathcal{R}(\mathbf{b}) = 2\,\mathrm{Re}\left[\mathbf{a}^H\mathbb{H}_c(\mathbf{h})\mathbf{b} + \mathbf{a}^H\bar{\mathbb{H}}_c(\mathbf{h})\mathbf{b}^*\right],$$

where $\mathbf{a}$ and $\mathbf{b}$ are vectors of appropriate dimensions. At true channel and applying the weak law of large numbers, we can see that [23]

$$\begin{aligned}\mathbf{a}^H\mathbb{H}_c\left(\bar{\mathbf{h}}\right)\mathbf{b} - \left(\mathbf{a}^H\mathbf{\Theta}^H\left(\frac{1}{\widetilde{\sigma}^2}\hat{\mathbf{P}} + \frac{1}{\widetilde{\sigma}_w^2}\hat{\mathbf{P}}^{\perp}\right)\mathbf{\Theta}\mathbf{b}\right) &\to 0\\ \mathbf{a}^H\bar{\mathbb{H}}_c\left(\bar{\mathbf{h}}\right)\mathbf{b} &\to 0\end{aligned} \tag{22}$$

with probability one, where $\widetilde{\sigma}^2 = \frac{1}{\hat{N}_{\mathrm{imp}}}\mathrm{tr}\left[\hat{\mathbf{P}}\bar{\mathbf{C}}_n\right]$ and $\widetilde{\sigma}_w^2 = \frac{1}{N-\hat{N}_{\mathrm{imp}}}\mathrm{tr}\left[\hat{\mathbf{P}}^{\perp}\bar{\mathbf{C}}_n\right]$.

Exploiting Eqs. (22) and (12) we can conclude that

$$\mathbb{H}\zeta_{RML}\left(\bar{\mathbf{h}}\right) - 2\begin{bmatrix}\mathrm{Re}\left[\mathbf{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\mathbf{\Theta}\right] & -\mathrm{Im}\left[\mathbf{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\mathbf{\Theta}\right]\\ \mathrm{Im}\left[\mathbf{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\mathbf{\Theta}\right] & \mathrm{Re}\left[\mathbf{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\mathbf{\Theta}\right]\end{bmatrix} \to \mathbf{0}, \tag{23}$$

where $\widetilde{\mathbf{C}}_n^{-1} = \frac{1}{\widetilde{\sigma}^2}\hat{\mathbf{P}} + \frac{1}{\widetilde{\sigma}_w^2}\hat{\mathbf{P}}^{\perp}$, almost surely.

4.2.2. Asymptotic distribution of the gradient of the cost function

Similar to the Hessian, for the evaluation of the asymptotic distribution of the cost function gradient, we exploit the identity that establishes

$$\mathcal{R}^T(\mathbf{a})\nabla\zeta_{RML}(\mathbf{h}) = 2\,\mathrm{Re}\left[\mathbf{a}^H\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*}\right], \tag{24}$$

for a complex vector $\mathbf{a}$ of appropriate dimension. Furthermore, from Eq. (14), upon applying strong law of large numbers [23]

$$\left.\frac{\bar{\mathbf{y}}^H\hat{\mathbf{P}}\bar{\mathbf{y}}}{\hat{N}_{\mathrm{imp}}}\right|_{\mathbf{h}=\bar{\mathbf{h}}} \to \widetilde{\sigma}^2 \quad \text{and} \quad \left.\frac{\bar{\mathbf{y}}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}}}{N-\hat{N}_{\mathrm{imp}}}\right|_{\mathbf{h}=\bar{\mathbf{h}}} \to \widetilde{\sigma}_w^2$$

almost surely. This further implies that

$$\sqrt{N}\mathbf{a}^H\left.\frac{\partial\zeta_{RML}(\mathbf{h})}{\partial\mathbf{h}^*}\right|_{\mathbf{h}=\bar{\mathbf{h}}} = -\frac{1}{\widetilde{\sigma}^2}\mathbf{a}^H\boldsymbol{\Theta}^H\hat{\mathbf{P}}\bar{\mathbf{y}} - \frac{1}{\widetilde{\sigma}_w^2}\mathbf{a}^H\boldsymbol{\Theta}^H\hat{\mathbf{P}}^\perp\bar{\mathbf{y}} + o_p(1). \tag{25}$$

4.2.3. Asymptotic variance of the random ML estimator

Plugging the values of the Hessian and the gradient of the cost function evaluated at $\mathbf{h}=\bar{\mathbf{h}}$ into Eq. (21), the channel estimation error of the random ML estimator can hence be expressed as

$$\sqrt{N}\left(\hat{\mathbf{h}}_{RML}-\bar{\mathbf{h}}\right) = \left(\boldsymbol{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\right)^{-1}\boldsymbol{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\bar{\mathbf{n}} + o_p(1). \tag{26}$$

Furthermore, the estimation error is asymptotically equivalent in law to a Gaussian random variable with zero mean and covariance being equal to

$$\mathbf{C}_{RML} = \left(\boldsymbol{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\right)^{-1}\boldsymbol{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\bar{\mathbf{C}}_n\widetilde{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\left(\boldsymbol{\Theta}^H\widetilde{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\right)^{-1}. \tag{27}$$

5. Performance assessment

In this section we analyze and compare the performance of both estimators in varied scenarios by exploiting the derived covariances in Eqs. (18) and (27). To begin with, we first consider an ideal scenario where the support of the IN is perfectly estimated and characterize the variance of both channel estimators. In the second assessment, we consider scenarios where the support of the IN is estimated with error and determine the variance of the channel estimation error for the proposed channel estimators.

5.1. Correct estimation of the IN support

In an ideal situation, when the estimated support and the true support of the IN match, then $\bar{\mathbf{P}}=\hat{\mathbf{P}}$ and $\bar{\mathbf{P}}^\perp=\hat{\mathbf{P}}^\perp$. In such a situation $\widetilde{\sigma}^2=\bar{\sigma}^2$, $\widetilde{\sigma}_w^2=\bar{\sigma}_w^2$ and $\widetilde{\mathbf{C}}_n=\bar{\mathbf{C}}_n=\hat{\mathbf{C}}_n$.

Hence, the variances of the channel estimation errors of the two estimators when the support of the IN is accurately estimated, can be expressed as:

$$\mathbf{C}_{DML} = \bar{\sigma}_w^2\left(\boldsymbol{\Theta}^H\bar{\mathbf{P}}^\perp\boldsymbol{\Theta}\right)^{-1} \tag{28}$$

and

$$\mathbf{C}_{RML} = \left(\boldsymbol{\Theta}^H\bar{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\right)^{-1}. \tag{29}$$

Recalling that $\bar{\mathbf{C}}_n^{-1}=\bar{\sigma}_w^{-2}\bar{\mathbf{P}}^\perp+\bar{\sigma}^{-2}\bar{\mathbf{P}}$, we see that $\bar{\mathbf{C}}_n^{-1}\geq\bar{\sigma}_w^{-2}\bar{\mathbf{P}}^\perp$, implying

$$\mathbf{C}_{RML} \leq \mathbf{C}_{DML}. \tag{30}$$

From Eq. (30) we can conclude that, provided that the support of the IN samples are correctly estimated, the random ML channel estimator outperforms the performance of the deterministic ML estimator as its variance is lower. The superior performance of the random ML estimator is reasonable as it exploits all samples in the received signal to determine the CIR coefficients, whereas the deterministic ML estimator only uses those samples that are not contaminated by the IN. However, when the IN power is very high than the background noise power, then $\bar{\mathbf{C}}_n^{-1}\approx\bar{\sigma}_w^{-2}\bar{\mathbf{P}}^\perp$ and, hence, the variances of both estimators converge to approximately similar values such that $\mathbf{C}_{RML}\approx\mathbf{C}_{DML}$.

Furthermore,

$$\mathbf{C}_{DML} = \bar{\sigma}_w^2\left(\boldsymbol{\Theta}^H\bar{\mathbf{P}}^\perp\boldsymbol{\Theta}\right)^{-1} = \mathbf{CRB}_{DML}$$

and

$$\mathbf{C}_{RML} = \left(\boldsymbol{\Theta}^H\bar{\mathbf{C}}_n^{-1}\boldsymbol{\Theta}\right)^{-1} = \mathbf{CRB}_{RML}$$

are the CRBs of the deterministic and the random ML estimators respectively.

5.2. Incorrect estimation of the IN support

In practice, it might happen that the support of the IN is estimated with error. In such a situation $\hat{\mathbf{P}} \neq \bar{\mathbf{P}}$ and $\hat{\mathbf{P}}^{\perp} \neq \bar{\mathbf{P}}^{\perp}$ further implying that $\hat{\mathbf{C}}_n \neq \bar{\mathbf{C}}_n$. In a realistic scenario, a mismatch between the estimated and the true noise covariance matrices typically takes place when the support of the IN samples is either underestimated or overestimated.

In order to investigate the performance of both estimators in underestimated and overestimated scenarios, first we consider a situation when the IN support is underestimated and formulate the estimation error variances for both approaches. Following the underestimated support scenario, we then characterize the channel estimation errors of both estimators considering a situation when the supports of the IN are overestimated.

5.2.1. Underestimated IN support

In this scenario we assume that, among all the IN affected samples, only a subset is identified. The estimated support of the IN is therefore a smaller subset of all the truly contaminated samples. Under this assumption, we see that $\bar{N}_{\text{imp}} > \hat{N}_{\text{imp}}$, where $\bar{N}_{\text{imp}}$ is the true number of IN samples, and $\bar{\mathbf{P}} - \hat{\mathbf{P}} > 0$, such that $\hat{\mathbf{P}}\bar{\mathbf{P}} = \hat{\mathbf{P}}$, $\hat{\mathbf{P}}\bar{\mathbf{P}}^{\perp} = 0$, $\tilde{\sigma}^2 = \bar{\sigma}^2$ and

$$\tilde{\sigma}_w^2 = \bar{\sigma}^2 \frac{\bar{N}_{\text{imp}} - \hat{N}_{\text{imp}}}{N - \hat{N}_{\text{imp}}} + \bar{\sigma}_w^2 \frac{N - \bar{N}_{\text{imp}}}{N - \hat{N}_{\text{imp}}}.$$

The variance of the deterministic ML estimator, under above assumptions, can hence be expressed as:

$$\begin{aligned}\mathbf{C}_{DML} = {} & \bar{\sigma}_w^2 \left(\boldsymbol{\Theta}^H \hat{\mathbf{P}}^{\perp} \boldsymbol{\Theta}\right)^{-1} \\ & + \left(\bar{\sigma}^2 - \bar{\sigma}_w^2\right) \left(\boldsymbol{\Theta}^H \hat{\mathbf{P}}^{\perp} \boldsymbol{\Theta}\right)^{-1} \boldsymbol{\Theta}^H \left(\bar{\mathbf{P}} - \hat{\mathbf{P}}\right) \boldsymbol{\Theta} \left(\boldsymbol{\Theta}^H \hat{\mathbf{P}}^{\perp} \boldsymbol{\Theta}\right)^{-1}. \end{aligned} \tag{31}$$

Similarly, the variance of the random ML estimator takes the form:

$$\begin{aligned}\mathbf{C}_{RML} = {} & \left(\boldsymbol{\Theta}^H \tilde{\mathbf{C}}_n^{-1} \boldsymbol{\Theta}\right)^{-1} \boldsymbol{\Theta}^H \\ & \times \left(\frac{1}{\tilde{\sigma}^2} \hat{\mathbf{P}} + \frac{\bar{\sigma}^2}{\tilde{\sigma}_w^4} \left(\bar{\mathbf{P}} - \hat{\mathbf{P}}\right) + \frac{\bar{\sigma}_w^2}{\tilde{\sigma}_w^4} \bar{\mathbf{P}}^{\perp} \right) \\ & \times \boldsymbol{\Theta} \left(\boldsymbol{\Theta}^H \tilde{\mathbf{C}}_n^{-1} \boldsymbol{\Theta}\right)^{-1}, \end{aligned} \tag{32}$$

where we have used the fact that $\tilde{\mathbf{C}}_n^{-1} \bar{\mathbf{C}}_n \tilde{\mathbf{C}}_n^{-1}$ in Eq. (27) is equivalent to $\left(\frac{1}{\tilde{\sigma}^2} \hat{\mathbf{P}} + \frac{\bar{\sigma}^2}{\tilde{\sigma}_w^4} (\bar{\mathbf{P}} - \hat{\mathbf{P}}) + \frac{\bar{\sigma}_w^2}{\tilde{\sigma}_w^4} \bar{\mathbf{P}}^{\perp} \right)$ when the support of the IN is underestimated.

5.2.2. Overestimated IN support

In this scenario, we assume that the estimated set of IN contaminated samples is larger than the true one and also that all the truly contaminated samples are retained in the set of estimated support.

Under these two assumptions, we see that $\hat{N}_{\text{imp}} > \bar{N}_{\text{imp}}$, $\hat{\mathbf{P}} = \bar{\mathbf{P}} + \mathbf{V}\mathbf{V}^H$ for some selection matrix $\mathbf{V}$ with dimensions $N \times (\hat{N}_{\text{imp}} - \bar{N}_{\text{imp}})$, resulting into $\bar{\mathbf{P}}\mathbf{V} = \mathbf{0}$ (such that $\bar{\mathbf{P}}^{\perp}\mathbf{V} = \mathbf{V}$) implying $\hat{\mathbf{P}}\bar{\mathbf{P}} = \bar{\mathbf{P}}$, $\hat{\mathbf{P}}^{\perp}\bar{\mathbf{P}} = \mathbf{0}$, $\tilde{\sigma}_w^2 = \bar{\sigma}_w^2$ and

$$\tilde{\sigma}^2 = \bar{\sigma}^2 \frac{\bar{N}_{\text{imp}}}{\hat{N}_{\text{imp}}} + \bar{\sigma}_w^2 \frac{\hat{N}_{\text{imp}} - \bar{N}_{\text{imp}}}{\hat{N}_{\text{imp}}}.$$

Upon evaluating the variance of the deterministic ML estimator, the resulting variance can be readily expressed as:

$$\mathbf{C}_{DML} = \bar{\sigma}_w^2 \left(\boldsymbol{\Theta}^H \hat{\mathbf{P}}^{\perp} \boldsymbol{\Theta}\right)^{-1}. \tag{33}$$

Similarly, the variance of the random ML estimator can be expressed as:

$$\begin{aligned}\mathbf{C}_{RML} = {} & \left(\boldsymbol{\Theta}^H \tilde{\mathbf{C}}_n^{-1} \boldsymbol{\Theta}\right)^{-1} \boldsymbol{\Theta}^H \\ & \times \left(\frac{\bar{\sigma}^2}{\tilde{\sigma}^4} \bar{\mathbf{P}} + \frac{\bar{\sigma}_w^2}{\tilde{\sigma}^4} \left(\hat{\mathbf{P}} - \bar{\mathbf{P}}\right) + \frac{1}{\tilde{\sigma}_w^2} \hat{\mathbf{P}}^{\perp} \right) \\ & \times \boldsymbol{\Theta} \left(\boldsymbol{\Theta}^H \tilde{\mathbf{C}}_n^{-1} \boldsymbol{\Theta}\right)^{-1}, \end{aligned} \tag{34}$$

where we have used the fact that taking into account when the IN support is over estimated, $\tilde{\mathbf{C}}_n^{-1} \bar{\mathbf{C}}_n \tilde{\mathbf{C}}_n^{-1}$ in Eq. (27) is equivalent to $(\frac{\bar{\sigma}^2}{\tilde{\sigma}^4} \bar{\mathbf{P}} + \frac{\bar{\sigma}_w^2}{\tilde{\sigma}^4} (\hat{\mathbf{P}} - \bar{\mathbf{P}}) + \frac{1}{\tilde{\sigma}_w^2} \hat{\mathbf{P}}^{\perp})$.

Remark 1. After evaluating the variances of the proposed estimators by segregating the scenarios when the estimated noise covariance does not match the true one, identifying a better performing channel estimator between the two proposed ones, as done for the ideal scenario where the support of the IN is accurately estimated, appears to be a challenging task. Therefore, in order to validate the performances of the proposed estimators in IN support underestimated and overestimated scenarios, we resort to channel estimation simulations for both estimators and validate their performances numerically.

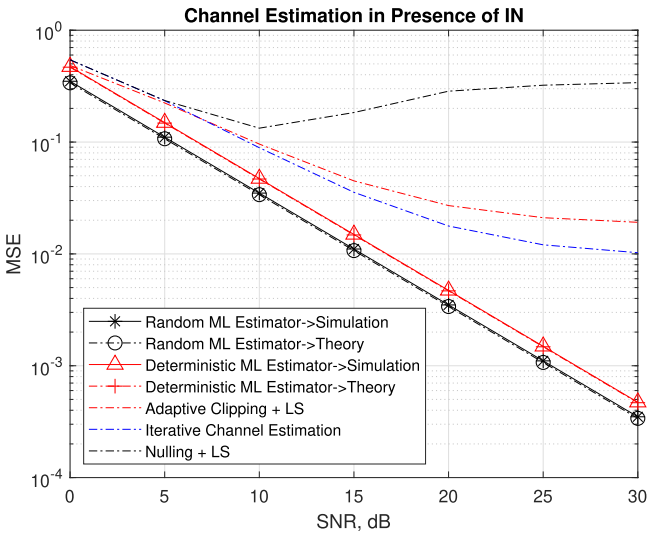

Fig. 3. Performance of both estimators when impulsive to background noise power ratio is fixed to 10 dB and signal to noise power ratio is varied from 0 dB to 30 dB.

6. Numerical validation

In this section, we numerically evaluate the performance of the proposed channel estimators in typical PLC scenarios using MATLAB. For the purpose of simulation, we consider an OFDM based PLC system with $N = 256$ subcarriers. Among the 256 subcarriers, only $N_u = 252$ subcarriers are used for data transmission. Each sub carrier conveys complex symbols that are randomly drawn from an uncoded QPSK modulation. To depict a realistic power line channel, the frequency selective multipath channel model proposed in IEEE 1901.2 is considered for the numerical validation. The values of all the channel parameters are adopted as proposed in the standard [8]. Furthermore, the occurrence of IN is defined according to the Bernoulli-Gaussian noise model [11]. Following this model, the occurrence of IN samples is assumed to be Bernoulli distributed and the amplitude of noise samples is assumed to be Gaussian distributed with zero mean and $\tilde{\sigma}_i^2$ variance. A sample of such IN is hence denoted as:

$$[\mathbf{i}_p]_k = b_k g_k,$$

where k denotes the k^{th} sample of IN, b_k denotes the Bernoulli distributed random variable and g_k denotes the Gaussian distributed random variable.

In order to compare the performance of the proposed channel estimators with the channel estimation strategies in the literature, we also run the numerical validations of the nulling + LS, adaptive clipping + LS and iterative channel estimation algorithms [11,24,25]. To perform simulations, we set the threshold for the stopping criterion for the random ML estimator to $t = 10^{-5}$ in Eq. (10) and the Newton-Raphson method is initialized by $\mathbf{h}_0 = \left[(\Theta^H \hat{\mathbf{C}}_n^{-1} \Theta)^{-1} \Theta^H \hat{\mathbf{C}}_n^{-1} \mathbf{y}\right] / \sqrt{N}$ as the initial guess of $\mathbf{h}$.

Fig. 3 shows mean-squared error (MSE) of the proposed channel estimators and the conventional channel estimators in a scenario where INR is fixed to 10 dB and SNR is varied from 0 dB to 30 dB. As shown, the conventional nulling + LS estimator yields highest MSE among the estimators under comparison and its performance tends to saturate in the region where SNR is higher than 20 dB. The clipping + LS scheme based on adaptive thresholding shows better performance than the nulling + LS estimator. However, this estimator also shows saturating behavior at high SNR region. Similar type of saturation behavior is also shown by the iterative channel estimator but yields lower MSE. Different to these, the proposed channel estimator that exploits the estimated IN treating it as a random quantity not only outperforms the performance of the deterministic ML estimator but also yields lower MSE in comparison to the conventional channel estimation strategies. The superior performance of the random ML estimator holds true at different SNR values and provides significant gain in terms of low MSE for the channel estimation. The better performance of the random ML estimator comes from the fact that it exploits all the samples in the received signal to approximate the coefficients of CIR. In contrast, the deterministic ML estimator, which exploits only the non-contaminated samples in the received signal, results in a higher variance than the random approach. Nonetheless, both estimators, under accurate IN support estimation, fairly reach their corresponding CRBs as shown in Fig. 3. Since a significant amount of gain in channel estimation can be achieved by the proposed estimators, in the upcoming numerical validation results we do not consider showing results of conventional channel estimation strategies and focus more on analyzing the behavior of the proposed channel estimators in typical impulsive PLC scenarios.

In a situation when the SNR is fixed to 20 dB and INR is varied from 10 dB to 40 dB, the performance of both estimators tends to converge to a saturated variance in the region where the IN bears very high power in comparison to the background

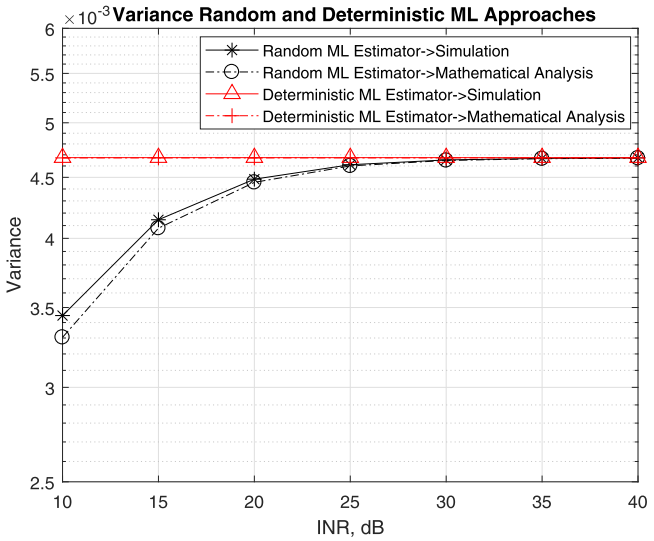

Fig. 4. Performance of both estimators when the signal to background noise power ratio is fixed to 20 dB and the IN to background noise power is varied from 10 dB to 40 dB.

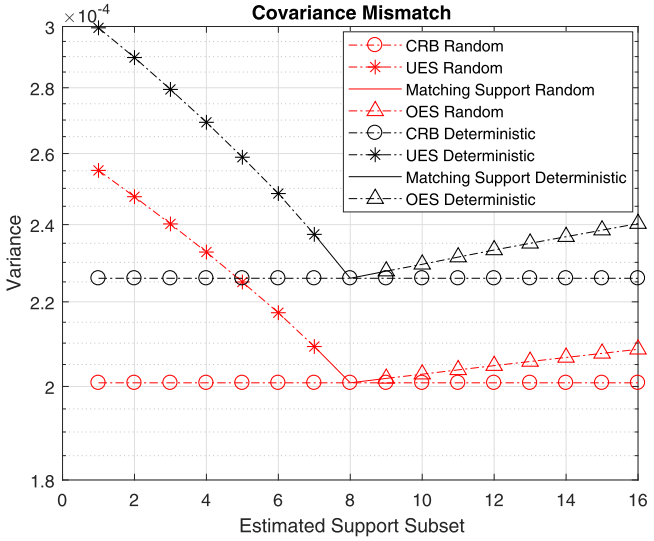

Fig. 5. Performance of proposed estimators when the impulsive to background noise power ratio is fixed to 10 dB and the signal to background noise power is fixed to 20 dB under noise covariance mismatch scenarios. UES and OES in the legend respectively denote IN support underestimated and overestimated scenarios.

noise, as reported in Fig. 4. However, at low INR regions, the random ML estimator significantly outperforms deterministic ML estimator.

Referring to the scenarios defined in Section 5.2, numerical validation of the performance of both approaches is done considering situations when the support of the IN is estimated with error. Regardless of whether the support of the IN is underestimated and overestimated, the random ML estimator is found to outperform the deterministic ML channel estimator, as shown in Fig. 5. In a situation where the support of the IN is underestimated, the channel estimator following the deterministic approach cannot distinguish all the IN corrupted samples. As a consequence, the performance of the deterministic ML estimator degrades resulting into having higher channel estimation error variance as compared to the random ML estimator. Similarly when the IN samples supports are over estimated, the number of samples available for channel estimation to the deterministic ML estimator is reduced and hence higher variance is perceived by the estimator. In contrast, the random ML estimator still has the opportunity to eliminate all the IN contaminated samples and approximate the impulse response of the channel, resulting into lower variance than the deterministic ML channel estimator. Nevertheless, both estimators reach their respective CRBs when the estimated support and true support coincide. These effects are well demonstrated in the simulation result reported in Fig. 5. The x-axis in Fig. 5 is scaled to denote subsets of true IN support

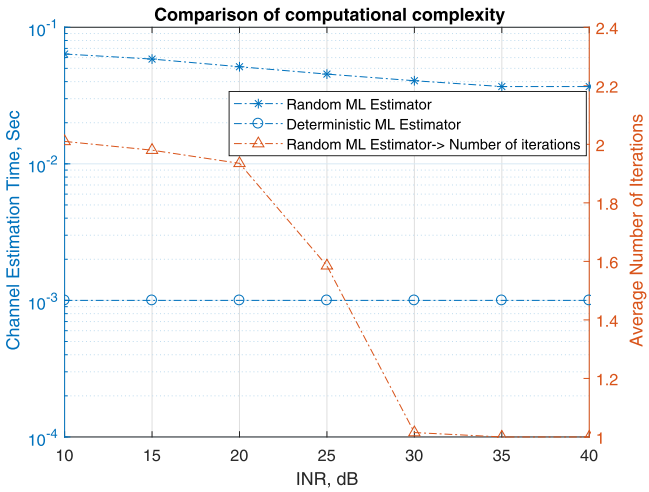

Fig. 6. Computational efficiency in terms of time taken by the proposed estimators to estimate the channel in scenarios where the signal to background noise power ratio is fixed at 20 dB and the impulsive to background noise power ratio is varied from 10 dB to 40 dB.

set when they are underestimated *(extreme at the left)* and ranging up to the situation where the noise support is overestimated *(extreme at the right)*. The mid-point, denoted by covariance matching point in the figure, defines the situation when the estimated support and the true support of IN sample coincide.

Despite the fact that the random ML estimator yields an estimate of the channel with more precision, the cost in terms of computational efficiency however needs to be paid. As shown in Fig. 6, where we evaluate the computational time taken by both estimators to estimate the channel, the deterministic ML estimator is computationally inexpensive and is timely consistent on estimating the channel on varied INR scenarios (fixed SNR). The random ML estimator however shows efficient channel estimation capability only at high INR regions and shows computational inefficiency in low INR regions. The computational load in random ML channel estimation is high because, in each iteration for the CIR estimation, the Hessian matrix and gradient vector need to be calculated along with the matrix inversion and matrix multiplication operations.

7. Conclusion

Two ML channel estimators for PLC systems have been proposed in this paper. Both estimators exploit the estimated IN samples to robustly determine the coefficients of CIR in an impulsive environment and differ based on the way they treat the estimated IN while estimating the channel.

The estimation errors of both estimators have been evaluated and analytically characterized. Performances of both estimators have been assessed and validated numerically. As verified by the simulation results, both estimators reach their respective Cramer-Rao bounds when the support of IN is estimated accurately. Furthermore, in all typical PLC scenarios, the random ML estimator has been found to outperform the performance of deterministic ML estimator. Nevertheless the deterministic ML estimator is computationally lighter than the random approach and offers a consistent channel estimation strategy in situations where higher computational complexity cannot be afforded.

Acknowledgment

The work presented in this paper is supported in part by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 607774 ADVANTAGE and by the Generalitat de Catalunya under grants 2014 SGR 1551, 2014 SGR 1567 and 2017 SGR 891.

Appendix A. Consistency of the likelihood function for the random approach based estimator

In order to check the consistency of the random ML estimator, we assume that the sample size N grows to infinity and the number of subcarriers bearing pilot symbols N_u along with the number of IN samples N_{imp} are a function of N such that

$$0 < \liminf_N \frac{N_{\text{imp}}}{N} \leq \limsup_N \frac{N_{\text{imp}}}{N} < +\infty,$$

$$0 < \liminf_N \frac{N_u}{N} \leq \limsup_N \frac{N_u}{N} < 1.$$

Furthermore, when N grows large, we assume that the diagonal values of $\mathbf{\Lambda}_p^H \mathbf{\Lambda}_p$ are contained in a fixed interval of the positive real axis.

Recalling from Eq. (9),

$$\zeta_{RML}(\mathbf{h}) = \frac{N_{\text{imp}}}{N} \log\left(\frac{1}{N_{\text{imp}}} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h}\right)^H \mathbf{P} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h}\right) \right) + \frac{N - N_{\text{imp}}}{N} \log\left(\frac{1}{N - N_{\text{imp}}} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h}\right)^H \mathbf{P}^{\perp} \left(\mathbf{y} - \sqrt{N}\mathbf{\Theta}\mathbf{h}\right) \right),$$

or, equivalently,

$$\begin{aligned}\zeta_{RML}(\mathbf{h}) = \frac{N_{\text{imp}}}{N} \log\Bigg[\frac{N}{N_{\text{imp}}} \left(\mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}\right) + \frac{\mathbf{n}}{\sqrt{N}} \right)^H \\ \times \mathbf{P} \left(\mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}\right) + \frac{\mathbf{n}}{\sqrt{N}} \right) \Bigg] \\ + \frac{N - N_{\text{imp}}}{N} \log\Bigg[\frac{N}{N - N_{\text{imp}}} \left(\mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}\right) + \frac{\mathbf{n}}{\sqrt{N}} \right)^H \\ \times \mathbf{P}^{\perp} \left(\mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}\right) + \frac{\mathbf{n}}{\sqrt{N}} \right) \Bigg].\end{aligned}$$

In order to prove point wise consistency of the random approach based cost function, we establish the following proposition.

Proposition 1. *Under above conditions, we have that*

$$\zeta_{RML}(\mathbf{h}) - \bar{\zeta}_{RML}(\mathbf{h}) \to 0$$

almost surely, where

$$\bar{\zeta}_{RML}(\mathbf{h}) = \frac{N_{imp}}{N} \log\left(\frac{N}{N_{imp}} \left(\bar{\mathbf{h}} - \mathbf{h}\right)^H \mathbf{\Theta}^H \mathbf{P} \mathbf{\Theta} \left(\bar{\mathbf{h}} - \mathbf{h}\right) + \tilde{\sigma}^2 \right) + \frac{N - N_{imp}}{N} \log\left(\frac{N}{N - N_{imp}} \left(\bar{\mathbf{h}} - \mathbf{h}\right)^H \mathbf{\Theta}^H \mathbf{P}^{\perp} \mathbf{\Theta} \left(\bar{\mathbf{h}} - \mathbf{h}\right) + \tilde{\sigma}_w^2 \right),$$

taking into account

$$\tilde{\sigma}^2 = \frac{1}{N_{imp}} \text{tr}\left[\mathbf{P}\bar{\mathbf{C}}_n\right] \quad \textit{and} \quad \tilde{\sigma}_w^2 = \frac{1}{N - N_{imp}} \text{tr}\left[\mathbf{P}^{\perp}\bar{\mathbf{C}}_n\right].$$

Before sketching a proof of the above result, it is interesting to note that the random method appears to be consistent even if the true noise covariance $\bar{\mathbf{C}}_n$ is completely different from the one assumed in the model (note that this is a property that is shared with the deterministic ML approach).

Proof. Observe that we can write

$$\zeta_{RML}(\mathbf{h}) - \bar{\zeta}_{RML}(\mathbf{h}) = \frac{N_{\text{imp}}}{N} \log\left(\frac{\left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}}\right)^H \mathbf{P} \left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}}\right)}{\mathbf{m}^H \mathbf{P} \mathbf{m} + \frac{N_{\text{imp}}}{N} \tilde{\sigma}^2} \right) + \frac{N - N_{\text{imp}}}{N} \log\left(\frac{\left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}}\right)^H \mathbf{P}^{\perp} \left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}}\right)}{\mathbf{m}^H \mathbf{P}^{\perp} \mathbf{m} + \frac{N - N_{\text{imp}}}{N} \tilde{\sigma}_w^2} \right).$$

where we have denoted $\mathbf{m} = \mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}\right)$ for simplicity.

Applying the strong law of large numbers, it is trivial to observe that the two arguments of the logarithms converge to one, such that the preposition is verified. □

After establishing the point wise convergence in probability of the cost function, we now assess the consistency of the estimate of the channel impulse response. This does not follow from point wise convergence and must be extended to uniform convergence over a certain compact subset.

Let us fix a positive constant $C > 0$ large enough, and denote by $\mathcal{K} \subset \mathbb{C}^L$ the compact subset

$$\mathcal{K} = \left\{ \mathbf{h} \in \mathbb{C}^L : \limsup_N \left\| \bar{\mathbf{h}} - \mathbf{h} \right\|^2 \leq C \right\}.$$

The following proposition generalizes the consistency of the random ML cost function under the above assumptions, uniformly in $\mathcal{K}$.

Proposition 2. *Under the above assumptions,*

$$\sup_{\mathbf{h} \in \mathcal{K}} \left| \zeta_{RML}(\mathbf{h}) - \bar{\zeta}_{RML}(\mathbf{h}) \right| \to 0$$

almost surely.

Proof. Using the fact that $\log(1+x) < x$ for $x > -1$, the triangular inequality and the fact that $\mathbf{m}^H \mathbf{P} \mathbf{m} \geq \mathbf{0}$, $\mathbf{m}^H \mathbf{P}^\perp \mathbf{m} \geq \mathbf{0}$, we see that

$$\left| \zeta_{RML}(\mathbf{h}) - \bar{\zeta}_{RML}(\mathbf{h}) \right| \leq \frac{|\epsilon(\mathbf{m})|}{\widetilde{\sigma}^2} + \frac{\left|\epsilon^\perp(\mathbf{m})\right|}{\widetilde{\sigma}_w^2},$$

where

$$\epsilon(\mathbf{m}) = \left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}} \right)^H \mathbf{P} \left(\mathbf{m} + \frac{\mathbf{n}}{\sqrt{N}} \right) - \left(\mathbf{m}^H \mathbf{P} \mathbf{m} + \frac{N_{\text{imp}}}{N} \widetilde{\sigma}^2 \right)$$

and where $\epsilon^\perp(\mathbf{m})$ is equivalently defined by replacing $\mathbf{P}$ with $\mathbf{P}^\perp$. Note that

$$\widetilde{\sigma}^2 = \frac{1}{N_{\text{imp}}} \text{tr}\left[\mathbf{P} \bar{\mathbf{C}}_n \right] \geq \lambda_{\min}\left(\bar{\mathbf{C}}_n \right) > 0,$$

which implies that $\inf_N \widetilde{\sigma}^2 > 0$ because the eigenvalues of $\bar{\mathbf{C}}_n$ are bounded away from zero. A similar conclusion holds true for $\widetilde{\sigma}_w^2$. We therefore only need to show that $\sup_{\mathbf{h} \in \mathcal{K}} |\epsilon(\mathbf{m})| \to 0$ and $\sup_{\mathbf{h} \in \mathcal{K}} \left|\epsilon^\perp(\mathbf{m})\right| \to 0$ almost surely. Let us prove this for $\epsilon(\mathbf{m})$, the proof for $\epsilon^\perp(\mathbf{m})$ following the same arguments. Note that, by the triangular inequality,

$$|\epsilon(\mathbf{m})| \leq 2 \left| \frac{\mathbf{m}^H \mathbf{P} \mathbf{n}}{\sqrt{N}} \right| + \left| \frac{\mathbf{n}^H \mathbf{P} \mathbf{n}}{N} - \frac{N_{\text{imp}}}{N} \widetilde{\sigma}^2 \right|.$$

The second term does not depend on $\mathbf{m}$ and converges almost surely to zero by the strong law of the large numbers. It remains to show that the first term above, which will be denoted by $\epsilon_1(\mathbf{m})$ converges almost surely to zero. To see this, consider a network of vectors $\mathcal{H} = \{\mathbf{h}_k \in \mathcal{K}, k \in \mathcal{J}\}$ such that

$$\sup_{\mathbf{h} \in \mathcal{K}} \min_{k \in \mathcal{J}} \|\mathbf{h} - \mathbf{h}_k\| < \frac{1}{N}.$$

This can clearly be achieved for $|\mathcal{J}| = cN^{2L}$ with $c > 0$ being a positive constant independent of N. For a given $\mathbf{h}$, let k be an index (that depends on $\mathbf{h}$) such that $\|\mathbf{h} - \mathbf{h}_k\| < N^{-1}$, and denote $\mathbf{m}_k = \mathbf{\Theta}\left(\bar{\mathbf{h}} - \mathbf{h}_k\right)$. We can decompose

$$\epsilon_1(\mathbf{m}) = \epsilon_1(\mathbf{m}) - \epsilon_1(\mathbf{m}_k) + \epsilon_1(\mathbf{m}_k)$$

and observe that, by Cauchy-Schwarz,

$$\begin{aligned} |\epsilon_1(\mathbf{m}) - \epsilon_1(\mathbf{m}_k)|^2 &= \left| \frac{(\mathbf{m} - \mathbf{m}_k)^H \mathbf{P} \mathbf{n}}{\sqrt{N}} \right|^2 \\ &\leq \|\mathbf{m} - \mathbf{m}_k\|^2 \frac{\mathbf{n}^H \mathbf{P} \mathbf{n}}{N} \\ &\leq \lambda_{\max}\left(\mathbf{\Theta}^H \mathbf{\Theta} \right) \|\mathbf{h} - \mathbf{h}_k\|^2 \frac{\mathbf{n}^H \mathbf{P} \mathbf{n}}{N} \\ &\leq \frac{\lambda_{\max}\left(\mathbf{\Theta}^H \mathbf{\Theta} \right)}{N^2} \frac{\mathbf{n}^H \mathbf{P} \mathbf{n}}{N}, \end{aligned}$$

from where it follows that $\sup_{\mathbf{h} \in \mathcal{K}} |\epsilon_1(\mathbf{m}) - \epsilon_1(\mathbf{m}_k)| \to 0$ almost surely by the Borell-Cantelli lemma (note that $\sup_N \lambda_{\max}\left(\mathbf{\Theta}^H \mathbf{\Theta} \right) < \infty$ by assumption). On the other hand, we can also see that for any $\epsilon > 0$ and $r > 0$, we have

$$\begin{aligned} \mathbb{P}\left(\sup_{k \in \mathcal{J}} |\epsilon_1(\mathbf{m}_k)| > \epsilon \right) &\leq cN^{2L} \sup_{k \in \mathcal{J}} \mathbb{P}(|\epsilon_1(\mathbf{m}_k)| > \epsilon) \\ &\leq cN^{2L} \sup_{k \in \mathcal{J}} \frac{\mathbb{E}\left[|\epsilon_1(\mathbf{m}_k)|^{2r} \right]}{\epsilon^{2r}} \end{aligned}$$

$$\leq cN^{2L} \sup_{k\in\mathcal{J}} \frac{\mathbb{E}\left[\left|(\mathbf{m}_k)^H \mathbf{P}\mathbf{n}\right|^{2r}\right]}{\epsilon^{2r}N^r}$$

$$= cN^{2L}\left(\sup_{k\in\mathcal{J}} (\mathbf{m}_k)^H \mathbf{P}\mathbf{m}_k\right)^r \frac{\mathbb{E}\left[|n|^{2r}\right]}{\epsilon^{2r}N^r},$$

where n is a standard complex Gaussian random variable. By choosing $r > 2L + 1$ and applying the Borel-Cantelli lemma we obtain the desired result. □

Having established uniform convergence of the random ML cost function, consistency of the random ML channel estimates follows from the standard argument.

References

[1] Ancillotti E, Bruno R, Conti M. The role of communication systems in smart grids: architectures, technical solutions and research challenges. Comput Commun 2013;36(17):1665–97. doi:10.1016/j.comcom.2013.09.004. http://www.sciencedirect.com/science/article/pii/S0140366413002090

[2] Mouftah HT, Erol-Kantarci M, Rehmani MH. Transportation and power grid in smart cities: communication networks and services. John Wiley & Sons; 2018.

[3] Rehmani MH, Kantarci ME, Rachedi A, Radenkovic M, Reisslein M. IEEE access special section editorial smart grids: a hub of interdisciplinary research. IEEE Access 2015;3:3114–18. doi:10.1109/ACCESS.2016.2516158.

[4] Galli S, Scaglione A, Wang Z. For the grid and through the grid: the role of power line communications in the smart grid. Proc IEEE 2011;99(6):998–1027. doi:10.1109/JPROC.2011.2109670.

[5] Andreadou N, Pavlidou FN. PLC channel: impulsive noise modelling and its performance evaluation under different array coding schemes. IEEE Trans Power Deliv 2009;24(2):585–95. doi:10.1109/TPWRD.2008.2002958.

[6] Khan AA, Rehmani MH, Reisslein M. Cognitive radio for smart grids: survey of architectures, spectrum sensing mechanisms, and networking protocols. IEEE Commun Surv Tutor 2016;18(1):860–98. doi:10.1109/COMST.2015.2481722.

[7] Zimmermann M, Dostert K. A multipath model for the powerline channel. IEEE Trans Commun 2002;50(4):553–9. doi:10.1109/26.996069.

[8] IEEE Standards Association, IEEE Standard for Low-Frequency (less than 500 kHz) Narrowband Power Line Communications for Smart Grid Applications, Oct. 31, 2013, 269 pgs. doi:10.1109/IEEESTD.2013.6679210.

[9] Yang L, Ren G, Qiu Z. Novel noise reduction algorithm for LS channel estimation in OFDM system with frequency selective channels. In: IEEE International Conference on Communication Systems; 2010. p. 478–82. doi:10.1109/ICCS.2010.5686663.

[10] Torio P, Sanchez MG. Method to cancel impulsive noise from power-line communication systems by processing the information in the idle carriers. IEEE Trans Power Deliv 2012;27(4):2421–2. doi:10.1109/TPWRD.2012.2209776.

[11] Chien YR. Iterative channel estimation and impulsive noise mitigation algorithm for OFDM-based receivers with application to power-line communications. IEEE Trans Power Deliv 2015;30(6):2435–42. doi:10.1109/TPWRD.2015.2445925.

[12] Yu X, Lin P, He Z, Wu W. OFDM channel estimation with impulse noise cancellation. In: International Conference on Wireless Communications, Networking and Mobile Computing; 2007. p. 330–3. doi:10.1109/WICOM.2007.89.

[13] Yasui H, Ohno K, Itami M. A study on channel estimation using pilot symbols under impulsive PLC channel. In: 18th IEEE International Symposium on Power Line Communications and Its Applications; 2014. p. 322–7. doi:10.1109/ISPLC.2014.6812352.

[14] Ando H, Nakamura A, Ohno K, Itami M. A study on channel estimation under class A impulsive PLC channel. In: IEEE 17th International Symposium on Power Line Communications and Its Applications; 2013. p. 69–74. doi:10.1109/ISPLC.2013.6525827.

[15] Picorone AM, Amado LR, Ribeiro MV. Linear and periodically time-varying PLC channels estimation in the presence of impulsive noise. In: International Symposium on Power Line Communications and its Applications (ISPLC); 2010. p. 255–60. doi:10.1109/ISPLC.2010.5479915.

[16] Xu C, Chen X, Yu S. A modified time-domain channel estimation method for OFDM systems in impulsive noise environment. In: Fourth International Conference on Intelligent Control and Information Processing (ICICIP); 2013. p. 88–93. doi:10.1109/ICICIP.2013.6568046.

[17] Huang J, Wang P, Wan Q. Robust approach for channel estimation in power line communication. J Commun Netw 2012;14(3):237–42. doi:10.1109/JCN.2012.6253083.

[18] Shrestha D, Mestre X, Payar M. Maximum-likelihood channel estimation in presence of impulsive noise for PLC systems. In: IEEE Global Conference on Signal and Information Processing (GlobalSIP); 2016. p. 20–4. doi:10.1109/GlobalSIP.2016.7905795.

[19] Lin J, Nassar M, Evans BL. Impulsive noise mitigation in powerline communications using sparse bayesian learning. IEEE J Sel Areas Commun 2013;31(7):1172–83. doi:10.1109/JSAC.2013.130702.

[20] Lampe L. Bursty impulse noise detection by compressed sensing. In: IEEE International Symposium on Power Line Communications and Its Applications; 2011. p. 29–34. doi:10.1109/ISPLC.2011.5764411.

[21] Shrestha D, Mestre X, Payar M. Asynchronous impulsive noise mitigation based on subspace support estimation for PLC systems. In: International Symposium on Power Line Communications and its Applications (ISPLC); 2016. p. 1–6. doi:10.1109/ISPLC.2016.7476266.

[22] Ackleh AS, Allen EJ, Kearfott RB, Seshaiyer P. Classical and modern numerical analysis: theory, methods, and practice. CRC PRESS; 2009.

[23] Feller W. An introduction to probability theory and its applications. vol. 2. 2nd. edition. John Wiley & Sons; 1970.

[24] Papilaya VN, Vinck AJH. Investigation on a new combined impulsive noise mitigation scheme for OFDM transmission. In: 2013 IEEE 17th International Symposium on Power Line Communications and Its Applications; 2013. p. 86–91. doi:10.1109/ISPLC.2013.6525830.

[25] Ndo G, Siohan P, Hamon MH. Adaptive noise mitigation in impulsive environment: application to power-line communications. IEEE Trans Power Deliv 2010;25(2):647–56. doi:10.1109/TPWRD.2009.2035505.

Deep Shrestha received the Electrical Engineering degree from the Tampere University of Technology in 2014. He is currently a Ph.D. candidate at Signal Theory and Communications Department of Universitat Politécnica de Catalunya (UPC). His research interest includes, but not limited to, signal processing, power line communication systems, smart grids, and ultra-low latency communication systems.

Xavier Mestre received the Electrical Engineering, Mathematics and PhD degrees from the Universitat Politécnica de Catalunya (UPC) in 1997, 2011 and 2002 respectively. In January 2003 he joined the Centre Tecnológic de Telecomunicacions de Catalunya (CTTC), where he currently holds a position as a Senior Research Associate and head of the Advanced Signal and Information Processing Department. He is associate editor of the IEEE Transactions on Signal Processing (2008-11, 2015-).

Miquel Payaró received the Electrical Engineering and the PhD degrees from the Universitat Politécnica de Catalunya in 2002 and 2007, respectively. He is now a senior researcher and the Head of the Communications Technologies Division at the Centre Tecnológic de Telecomunicacions de Catalunya, which he joined in 2009. He is serving as Associate Editor of IEEE Transactions on Wireless Communications.