Robust Estimation of the Channel Parameters in the Presence of Transmitter and Receiver IQ Impairments

Alexandru FRUNZĂ, Vincent CHOQUEUSE, Pascal MOREL, and Stéphane AZOU

***Abstract*—In this paper, we propose a new approach for the estimation of the channel parameters under In-phase/ Quadrature (IQ) impairments, where the term channel refers to any transformation occurring between the transmitter and receiver IQ impairments. While most of the available strategies for channel estimation under IQ impairments are based on the Least-Squares (LS) approach, we propose a new estimator based on the Maximum Likelihood (ML) technique. We show that a simple estimator obtained from the ML gives better statistical performance than any LS-based technique while having a significant lower computational complexity. Mathematically, the proposed ML-based estimator is obtained from the minimization of the smallest eigenvalue of a particular 2 × 2 matrix. Simulations show the benefit of the proposed approach for the estimation of the Carrier Frequency Offset (CFO) and for the estimation of Finite Impulse Response (FIR) channel coefficients under both transmitter and receiver IQ impairments.**

***Index Terms*—Maximum Likelihood Estimation, Digital Communication, Channel Estimation, IQ impairments.**

I. Introduction

In a communication chain, many effects and impairments can increase the communications error rate. These include the effect of the propagation channel, the modulator impairments, the carrier frequency deviations, and the impact of the different noise sources. To avoid any performance loss, a low-cost solution relies on the use of digital compensation algorithms. Most compensation algorithms require knowledge of the channel parameters and system impairments. In practice, these parameters are usually estimated from the received signal during a preprocessing stage.

In literature, many algorithms have been proposed for the estimation of the channel parameters and system impairments. In the last decade, a particular focus has been placed on the estimation of the receiver In-phase/Quadrature (IQ) impairment since this distortion can significantly deteriorate the system performance, especially for Orthogonal Frequency-Division Multiplexing (OFDM) systems. The available techniques for receiver IQ imbalance parameter estimation can be divided into two categories: blind and trained-based algorithms. Several techniques for the blind estimation of the receiver IQ imbalance have been proposed in [1-6]. These techniques are mainly based on the statistical properties of the real and imaginary parts of the transmitted signal (circularity or statistical independence).

While blind techniques have the benefit of not reducing the available bandwidth, trained-based techniques usually lead to better statistical performance. The joint estimation of the IQ and systems impairments with trained-based techniques has been addressed in many studies. In particular, several algorithms for the estimation of the receiver IQ imbalance and channel Finite Impulse Response (FIR) coefficients have been proposed in [7-12]. Extensions to the case where the communication system is also corrupted by Phase Noise (PN), Carrier Frequency Offset (CFO) and/or Direct Current (DC) offset have also been described in [13-18].

In addition to the receiver IQ impairments, the performance of a communication system can be also impacted by the transmitter IQ impairments. In [19], it is shown that a small transmitter impairment can cause significant degradation of communication performance. To avoid such performance loss, some authors have proposed to jointly estimate the transmitter and receiver IQ imbalance. In particular, several techniques have addressed the estimation of the IQ imbalance parameters and FIR channel coefficients in OFDM or Generalized Frequency Division Multiplexing (GFDM) systems [20-23]. Joint estimation of the carrier frequency, timing offsets and IQ parameters have also been considered in [19], [24-26]. In most of these studies, the estimation of the channel and system parameters is carried out using a Least Squares (LS) approach. Although this technique has the advantage of being simple to implement, the LS estimator is usually statistically suboptimal. When the focus is on the statistical performance, a better estimator is given by the Maximum Likelihood (ML) technique. Regarding the estimation of the channel parameters under both transmitter and receiver IQ impairments, it has been pointed out in [26] that the Maximum Likelihood estimator can be difficult to derive since the noise statistic is also impacted by the receiver IQ impairments. Nevertheless, it is shown in [27] that the ML estimator of the IQ imbalance parameters can be derived in closed form for the particular case of an Additive White Gaussian Noise (AWGN) channel under a particular IQ

This work was supported by the Embassy of France in Romania (Institut Français de Roumanie) and the Romanian Ministry of National Defense.

A. FRUNZA is with the Military Technical Academy "Ferdinand I", C.E.C.T.I, 050141 Bucharest, Romania, and Lab-STICC, UMR CNRS 6285, ENIB, Technopole Brest-Iroise, CS 73862, 29238 Brest Cedex 3, France (e-mail: alexandru.frunza@mta.ro, frunza@enib.fr)

V. CHOQUEUSE is with Lab-STICC, UMR CNRS 6285, ENIB, Technopole Brest-Iroise, CS 73862, 29238 Brest Cedex 3, France (e-mail: vincent.choqueuse@enib.fr).

P. MOREL is with Lab-STICC, UMR CNRS 6285, ENIB, Technopole Brest-Iroise, CS 73862, 29238 Brest Cedex 3, France (e-mail: pascal.morel@enib.fr).

S. AZOU is with Lab-STICC, UMR CNRS 6285, ENIB, Technopole Brest-Iroise, CS 73862, 29238 Brest Cedex 3, France (e-mail: azou@enib.fr).

Digital Object Identifier 10.32754/JMT.2022.2.04

imbalance configuration[1]. Even if the estimator reported in [27] is very specific, this work suggests that simple expressions could be also obtained for more general settings.

In this paper, we focus on the ML estimation of the channel parameters under both transmitter and receiver IQ imbalance, where the term channel is used in a broad sense to describe any linear effect and impairment occurring between the transmitter and receiver IQ imbalance. To estimate the channel parameters in the presence of IQ impairments, we propose to treat the IQ impairments as nuisance parameters. As compared to other studies, the proposed approach is more general because it can be applied to a larger class of channel models. This paper is organized as follows. Section II describes the signal model and Section III reviews some commonly used LS strategies for the estimation of the channel parameters. In section IV, we derive the expression of the exact and approximate ML estimator and Section V gives the expression of the Cramér-Rao bounds. Finally, Section VI illustrates the benefits of the proposed technique for two estimation problems: the estimation of the CFO with a known FIR channel and the estimation of the unknown FIR channel coefficients without CFO.

II. Signal Model

In this paper, we consider the general single-input single-output (SISO) communication system presented in Figure 1. The transmitted and received signals are described by two complex-valued column vectors of size N denoted by $\mathbf{x} = [x[0], \cdots, x[N-1]]^T$ and $\mathbf{y} = [y[0], \cdots, y[N-1]]^T$.

On the transmitter side, the complex signal is affected by the presence of IQ impairments. Mathematically, the effect of the IQ impairments can be modeled as [1], [8]:

$$z[n] \quad = \mu_t x[n] + \nu_t x^*[n] \tag{1}$$

where $(\mu_t, \nu_t) \in \mathbb{C}^2$ corresponds to the IQ imbalance parameters.

The distorted signal $z[n]$ is then transmitted to a general linear channel with Gaussian noise. The received signal $\mathbf{r} = [r[0], r[1], \cdots, r[N-1]]^T$ can be expressed as:

$$\mathbf{r} \quad = \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{z} + \mathbf{w}, \tag{2}$$

where $\mathbf{z} = [z[0], z[1], \cdots, z[N-1]]^T$, $\mathbf{H}(\boldsymbol{\Omega}_h)$ is a $N \times N$ complex-valued matrix that models the deterministic part of the channel, and w corresponds to the additive noise. We assume that the deterministic part of the channel depends on L real-valued unknown parameters denoted by $\boldsymbol{\Omega}_h$. This general expression can encompass different kinds of channel models and system impairments[2]. Regarding the additive noise, we assume that $\mathbf{w}$ is a white circular Gaussian noise, distributed as $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$, where $\mathbf{I}_N$ corresponds to the $N \times N$ identity matrix.

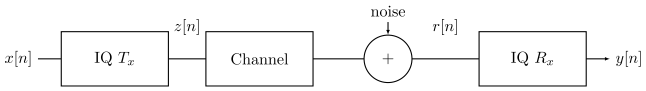


Figure 1. A SISO communication system with transmitter and receiver IQ impairments

[1]In [26], the IQ parameters are parameterized by only two complex coefficients.

[2]For example, this expression can model the influence of an FIR channel with DC offset, Carrier Frequency Offset, Carrier Phase noise, and the contribution of more specific linear effects such as Chromatic Dispersion for the case of optical communications.

At the receiver side, we consider that the signal is also corrupted by the presence of IQ imbalance. Therefore, the received signal is given by:

$$y[n] \quad = \mu_r r[n] + \nu_r r^*[n]. \tag{3}$$

Using (1), (2) and (3), the received signal can be expressed into a matrix form as:

$$\begin{aligned}\mathbf{y} \quad &= \mu_r\mu_t \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{x} + \mu_r\nu_t \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{x}^* + \\ &+ \nu_r\mu_t^* \mathbf{H}^*(\boldsymbol{\Omega}_h)\mathbf{x}^* + \nu_r\nu_t^* \mathbf{H}^*(\boldsymbol{\Omega}_h)\mathbf{x} + \\ &+ \mu_r \mathbf{w} + \nu_r \mathbf{w}^*.\end{aligned} \tag{4}$$

When all the IQ parameters are unknown, the signal model contains at least one parameter indetermination. Indeed, making the substitutions $\mu_t \to \alpha\mu_t$, $\mu_r \to \mu_r/\alpha$, $\nu_t \to \alpha\nu_t$, $\nu_r \to \nu_r/\alpha^*$ and $\sigma \to |\alpha|\sigma$ leave the signal model and the noise statistics unchanged. For this reason, without lack of generality, we assume that $\mu_r = 1$. Furthermore, for practical considerations, we also exclude the unrealistic case $|\nu_r| \geq 1$ from the estimation problem.

Using the assumption $\mu_r = 1$, the received signal can be expressed as follows:

$$\mathbf{y} \quad = \mathbf{s}(\boldsymbol{\Omega}_h) + \mathbf{b}, \tag{5}$$

where:

- $\mathbf{s}$ is a $N \times 1$ vector containing the deterministic part of the signal. This vector is defined as:

$$\mathbf{s}(\boldsymbol{\Omega}_h) \quad \triangleq \mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t + \nu_r \mathbf{F}^*(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t^*, \tag{6}$$

- with:

$$\mathbf{F}(\boldsymbol{\Omega}_h) \quad \triangleq \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{X}, \tag{7}$$

$$\boldsymbol{\theta}_t \quad \triangleq \begin{bmatrix}\mu_t \\ \nu_t\end{bmatrix}, \tag{8}$$

$$\mathbf{X} \quad \triangleq [\mathbf{x} \quad \mathbf{x}^*], \tag{9}$$

- $\mathbf{b} \triangleq \mathbf{w} + \nu_r \mathbf{w}^*$ is a $N \times 1$ vector containing the noise samples.

In addition to the complex-valued signal in (5), this paper also uses an augmented real-valued model. Let us introduce the $2N \times 1$ column vector $\tilde{\mathbf{y}} \triangleq [\Re e(\mathbf{y}^T), \Im m(\mathrm{y}^T)]^T$ obtained by concatenating the real and imaginary parts of the received signal $\mathbf{y}$. The augmented received vector can be expressed as:

$$\tilde{\mathbf{y}} \quad = \tilde{\mathbf{s}}(\boldsymbol{\Omega}_h) + \tilde{\mathbf{b}}, \tag{10}$$

where $\tilde{\mathbf{s}}(\boldsymbol{\Omega}_h) \triangleq [\Re e(\mathbf{s}(\boldsymbol{\Omega}_h)^T), \Im m(\mathbf{s}(\boldsymbol{\Omega}_h)^T)]^T$ is a $2N \times 1$ vector containing the augmented deterministic signal, and $\tilde{\mathbf{b}} = [\Re e(\mathbf{b}^T)), \Im m(\mathbf{b}^T))]^T$ corresponds to the augmented noise samples. Using (6), the vectors $\tilde{\mathbf{s}}(\boldsymbol{\Omega}_h)$ and $\tilde{\mathbf{b}}$ can be decomposed as:

$$\tilde{\mathbf{s}}(\boldsymbol{\Omega}_h) \quad = \begin{bmatrix}\Re e((1+\nu_r^*)\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t) \\ \Im m((1-\nu_r^*)\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t)\end{bmatrix}, \tag{11}$$

$$\tilde{\mathbf{b}} \quad = (\mathbf{M}(\nu_r) \otimes \mathbf{I}_N)\tilde{\mathbf{w}}, \tag{12}$$

Digital Object Identifier 10.32754/JMT.2022.2.04

where $\tilde{\mathbf{w}} = [\Re e(\mathbf{w}^T), \Im m(\mathbf{w}^T)]^T$, $\otimes$ corresponds to the Kronecker product, and:

$$\mathbf{M}(\nu_r) \triangleq \mathbf{I}_2 + \begin{bmatrix} \Re e(\nu_r) & \Im m(\nu_r) \\ \Im m(\nu_r) & -\Re e(\nu_r) \end{bmatrix}. \quad (13)$$

The goal of this paper is to estimate the unknown channel parameters $\mathbf{\Omega}_h$ from the received signal $\mathbf{y}$. As the received signal depends on the transmitter and receiver IQ imbalance parameters and on the noise variance, we propose to treat these additional unknowns as deterministic nuisance parameters. In this context, the number of (real-valued) parameters to be jointly estimated is equal to $L+7$. These parameters are given by:

$$\mathbf{\Omega} = [\mathbf{\Omega}_h^T, \mathbf{\Omega}_r^T, \mathbf{\Omega}_t^T, \sigma^2]^T \quad (14)$$

where:

- $\mathbf{\Omega}_h$ is a $L \times 1$ vector containing the real-valued channel parameters,
- $\mathbf{\Omega}_r = [\Re e(\nu_r), \Im m(\nu_r)]^T$ is a 2×1 vector containing the real and imaginary parts of the receiver IQ imbalance parameter ν_r,
- $\mathbf{\Omega}_t = [\Re e(\boldsymbol{\theta}_t^T), \Im m(\boldsymbol{\theta}_t^T)]^T$ is a 4×1 vector containing the real and imaginary parts of the transmitter IQ imbalance parameters,
- σ^2 is the noise variance.

III. Review of Least Squares-based Estimators

In the literature, several authors have proposed to use a LS approach for the estimation of the channel and IQ imbalance parameters. As the LS approach neglects the noise contribution, this strategy reduces the number of unknown parameters to $\mathbf{\Omega}_1 = [\mathbf{\Omega}_h^T, \mathbf{\Omega}_r^T, \mathbf{\Omega}_t^T]^T$. In this context, naive use of the LS estimator leads to the minimization of a $(L+6)$-dimensional cost function. In this section, we review some techniques to further reduce the number of dimensions of the LS cost function.

A. Naive LS

Using a naive LS approach, the estimator of the unknown parameters is given by:

$$\widehat{\mathbf{\Omega}}_1 = \arg\min_{\mathbf{\Omega}_1} \| \mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1) \|^2. \quad (15)$$

Due to the nonlinear dependence between $\mathbf{\Omega}_1$ and $\mathbf{s}(\mathbf{\Omega}_1)$, this technique is called the Nonlinear Least Squares Estimator (NLSE). The direct implementation of the NLSE requires the minimization of a $(L+6)$-dimensional cost function. Recently, several suboptimal approaches have been proposed based on cyclic minimization for the LS optimization problem [23], [27]. Nevertheless, these approaches are specific to particular channel models, and their extensions for more general settings are not trivial.

B. Separable LS estimator

To simplify the optimization problem, one possible solution is to replace some linear unknown variables by their estimate in the LS cost function. By splitting the real and imaginary parts of the received signal, we obtain the equivalent (real-valued) optimization problem:

$$\widehat{\mathbf{\Omega}}_1 = \arg\min_{\mathbf{\Omega}_1} \| \tilde{\mathbf{y}} - \mathbf{K}(\mathbf{\Omega}_h, \mathbf{\Omega}_r)\mathbf{\Omega}_t \|^2 \quad (16)$$

where $\mathbf{\Omega}_t = [\Re e(\boldsymbol{\theta}_e^T), \Im m(\boldsymbol{\theta}_e^T)]^T$ corresponds to the linear variables, and:

$$\mathbf{K}(\mathbf{\Omega}_h, \mathbf{\Omega}_r) \triangleq \begin{bmatrix} \Re e\big((1+\nu_r^*)\mathbf{F}(\mathbf{\Omega}_h)\big) & -\Im m\big((1+\nu_r^*)\mathbf{F}(\mathbf{\Omega}_h)\big) \\ \Im m\big((1-\nu_r^*)\mathbf{F}(\mathbf{\Omega}_h)\big) & \Re e\big((1-\nu_r^*)\mathbf{F}(\mathbf{\Omega}_h)\big) \end{bmatrix} \quad (17)$$

This optimization problem is a separable NLSE problem where the linear variables correspond to the real and imaginary parts of the transmitter IQ parameters $\boldsymbol{\theta}_t$. The estimator of the linear variables is: $\mathbf{\Omega}_t = \mathbf{K}^\dagger(\mathbf{\Omega}_h, \mathbf{\Omega}_r)\tilde{\mathbf{y}}$, with $\mathbf{K}^\dagger(\mathbf{\Omega}_h, \mathbf{\Omega}_r) = (\mathbf{K}^T(\mathbf{\Omega}_h, \mathbf{\Omega}_r)\mathbf{K}(\mathbf{\Omega}_h, \mathbf{\Omega}_r))^{-1}\mathbf{K}^T(\mathbf{\Omega}_h, \mathbf{\Omega}_r)$ the pseudo-inverse of $\mathbf{K}(\mathbf{\Omega}_h, \mathbf{\Omega}_r)$. By replacing this estimate into the cost function, we obtained the following concentrated LS cost function of $\mathbf{\Omega}_h$ and $\mathbf{\Omega}_r$ [28]:

$$\{\widehat{\mathbf{\Omega}}_h, \widehat{\mathbf{\Omega}}_r\} = \arg\min_{\mathbf{\Omega}_h, \mathbf{\Omega}_r} \tilde{\mathbf{y}}^T \mathbf{P}_{\mathbf{K}}^{\perp}(\mathbf{\Omega}_h, \mathbf{\Omega}_r)\tilde{\mathbf{y}} \quad (18)$$

where $\mathbf{P}_{\mathbf{K}}^{\perp}(\mathbf{\Omega}_h, \mathbf{\Omega}_r) = \mathbf{I}_{2N} - \mathbf{K}(\mathbf{\Omega}_h, \mathbf{\Omega}_r)\mathbf{K}^\dagger(\mathbf{\Omega}_h, \mathbf{\Omega}_r)$ is the orthogonal projector onto the null space of $\mathbf{K}^H(\mathbf{\Omega}_h, \mathbf{\Omega}_r)$. Note that this estimator leads to the same performance as the original naive LS approach and requires the minimization of a cost function along $(L+2 < L+6)$ dimensions.

C. LS with constraint relaxation

To fully decouple the estimation of the channel parameters from the estimation of the IQ parameters, another solution is to relax some constraints on the signal model. Using (6), the LS estimator can be reformulated as a constrained optimization problem as follows:

$$\widehat{\mathbf{\Omega}}_1 = \arg\min_{\mathbf{\Omega}_0} \| \mathbf{y} - \mathbf{G}(\mathbf{\Omega}_h)\boldsymbol{\theta} \|^2 \ \text{s.t}\ \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_t \\ \nu_r \boldsymbol{\theta}_t^* \end{bmatrix}, \quad (19)$$

where:

$$\mathbf{G}(\mathbf{\Omega}_h) \triangleq [\mathbf{F}(\mathbf{\Omega}_h), \mathbf{F}^*(\mathbf{\Omega}_h)]. \quad (20)$$

By relaxing the constraint on the structure of the vector $\boldsymbol{\theta}$ containing the transmitter and receiver IQ parameters, we obtain the following separable nonlinear LS problem:

$$\{\widehat{\mathbf{\Omega}}_h, \widehat{\boldsymbol{\theta}}\} = \arg\min_{\mathbf{\Omega}_h, \boldsymbol{\theta}} \| \mathbf{y} - \mathbf{G}(\mathbf{\Omega}_h)\boldsymbol{\theta} \|^2. \quad (21)$$

Using this strategy, the estimator of the channel parameters is simply given by:

$$\widehat{\mathbf{\Omega}}_h = \arg\min_{\mathbf{\Omega}_h} \mathbf{y}^H \mathbf{P}_{\mathbf{G}}^{\perp}(\mathbf{\Omega}_h)\mathbf{y}, \quad (22)$$

where $\mathbf{P}^{\perp}(\mathbf{\Omega}_h) = \mathrm{I}_N - \mathbf{G}(\mathbf{\Omega}_h)\mathbf{G}^\dagger(\mathbf{\Omega}_h)$ corresponds to the orthogonal projector onto the null space of $\mathbf{G}^H(\mathbf{\Omega}_h)$. This estimator requires the minimization of a L-dimensional cost function and fully decouples the estimation of the channel from the IQ impairments parameters. Nevertheless, due to the constraint relaxation, this approach is usually suboptimal. Note that the LS approach with constraint relaxation has been implicitly used in [26] for the estimation of the channel parameters.

When the noise distribution is unknown, the LS approach remains the most natural strategy for the estimation of the channel parameters. Moreover, when the noise is white, Gaussian, and circular after the receiver IQ imbalance, the Separable LS estimator corresponds to the Maximum Likelihood estimator and so, attains the Cramér Rao Bound at least asymptotically. Nevertheless, even if the noise can be reasonably assumed white, Gaussian and circular before the receiver IQ impairment, the circularity assumption generally does not hold after the receiver IQ imbalance [26], [27]. In this context, better estimators can be obtained by considering the non-circularity of the noise.

IV. Maximum-Likelihood based Estimator

In this section, we derive the Maximum Likelihood (ML) estimator of $\mathbf{\Omega}$ under the assumption of a circular white Gaussian noise occurring before the receiver IQ impairment. Under this assumption, $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_2)$ and the augmented received signal is distributed as follows:

$$\tilde{\mathbf{y}} \sim \mathcal{N}(\tilde{\mathbf{s}}(\mathbf{\Omega}_1), \mathbf{C}(\mathbf{\Omega}_2)), \quad (23)$$

where $\mathbf{\Omega}_1 = [\mathbf{\Omega}_h^T, \mathbf{\Omega}_r^T, \mathbf{\Omega}_t^T]^T$ is the vector containing the channel and IQ imbalance parameters and $\mathbf{\Omega}_2 = [\mathbf{\Omega}_r^T, \sigma^2]^T$ is a vector containing the receiver IQ parameter and the noise variance. The $2N \times 2N$ covariance matrix $\mathbf{C}(\mathbf{\Omega}_2)$ is given by:

$$\mathbf{C}(\mathbf{\Omega}_2) = \frac{\sigma^2}{2}(\mathbf{M}(\nu_r)\mathbf{M}^T(\nu_r) \otimes \mathbf{I}_N). \quad (24)$$

For the general linear model under Gaussian noise, the probability density function (pdf) of $\tilde{\mathbf{y}}$ is given by [29]:

$$p(\tilde{\mathbf{y}}; \mathbf{\Omega}) = \frac{1}{(2\pi)^{2N/2}\det^{1/2}(\mathbf{C}(\mathbf{\Omega}_2))} \times e^{-\frac{1}{2}(\tilde{\mathbf{y}}-\tilde{\mathbf{s}}(\mathbf{\Omega}_1))^T\mathbf{C}^{-1}(\mathbf{\Omega}_2)(\tilde{\mathbf{y}}-\tilde{\mathbf{s}}(\mathbf{\Omega}_1))}. \quad (25)$$

The determinant of the augmented covariance matrix $\mathbf{C}(\mathbf{\Omega}_2)$ is simply given by:

$$\det(\mathbf{C}(\mathbf{\Omega}_2)) = \left(\frac{\sigma^2\det(\mathbf{M}(\nu_r))}{2}\right)^{2N}, \quad (26)$$

with $\det(\mathbf{M}(\nu_r)) \triangleq 1 - |\nu_r|^2$. The inverse of the covariance matrix can also be expressed in closed form as follows:

$$\mathbf{C}^{-1}(\mathbf{\Omega}_2) = \frac{2(\mathbf{N}^T(\nu_r)\mathbf{N}(\nu_r) \otimes \mathbf{I}_N)}{\sigma^2\det^2(\mathbf{M}(\nu_r))}, \quad (27)$$

where:

$$\mathbf{N}(\nu_r) = \mathbf{I}_2 - \begin{bmatrix} \Re e(\nu_r) & \Im m(\nu_r) \\ \Im m(\nu_r) & -\Re e(\nu_r) \end{bmatrix}. \quad (28)$$

A. Estimation of the Noise variance

The Maximum Likelihood (ML) estimator of $\mathbf{\Omega}$ corresponds to the maximizer of $p(\tilde{\mathbf{y}}; \mathbf{\Omega})$ or equivalently to the maximiser of the log-likelihood function $\ln(p(\tilde{\mathbf{y}}; \mathbf{\Omega}))$. By maximizing the log-likelihood function with respect to σ^2, it can be checked that the ML estimator of the noise variance is given by:

$$\hat{\sigma}^2(\mathbf{\Omega}_1) = \frac{1}{N}\frac{\| (\mathbf{N}(\nu_r) \otimes \mathbf{I}_N)(\tilde{\mathbf{y}} - \tilde{\mathbf{s}}(\mathbf{\Omega}_1)) \|^2}{\det^2(\mathbf{M}(\nu_r))} = \frac{(1+|\nu_r|^2)r_{11}(\mathbf{\Omega}_1) - 2\Re e(\nu_r^* r_{12}(\mathbf{\Omega}_1))}{(1-|\nu_r|^2)^2}. \quad (29)$$

The parameters $r_{11}(\mathbf{\Omega}_1) \in \mathbb{R}^+$ and $r_{12}(\mathbf{\Omega}_1) \in \mathbb{C}$ are defined as:

$$r_{11}(\mathbf{\Omega}_1) = \frac{1}{N}(\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1))^H(\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1)) = \frac{1}{N}\| \mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1) \|^2, \quad (30)$$

$$r_{12}(\mathbf{\Omega}_1) = \frac{1}{N}(\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1))^T(\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_1)). \quad (31)$$

The two quantities $r_{11}(\mathbf{\Omega}_1)$ and $r_{12}(\mathbf{\Omega}_1)$ measure the variance and pseudo-variance of the received signal $\mathbf{y}$, respectively [30]. By replacing the noise variance by its estimate in the pdf of $\tilde{\mathbf{y}}$, we obtain:

$$p(\tilde{\mathbf{y}}; \mathbf{\Omega}_1, \hat{\sigma}^2) = \left(\frac{1}{\pi e\hat{\sigma}^2(\mathbf{\Omega}_1)\det(\mathbf{M}(\nu_r))}\right)^N. \quad (32)$$

The ML estimator of $\mathbf{\Omega}_1$ is given by the maximizer of $p(\tilde{\mathbf{y}}; \mathbf{\Omega}_1, \hat{\sigma}^2)$ or, similarly, by the maximizer of $\mathcal{L}(\mathbf{\Omega}_1) \triangleq (\pi e) \times (p(\tilde{\mathbf{y}}; \mathbf{\Omega}_1, \hat{\sigma}^2))^{1/N}$. Using the expression of $\hat{\sigma}^2(\mathbf{\Omega}_1)$ and the equality $\det(\mathbf{M}(\nu_r)) \triangleq 1 - |\nu_r|^2$, we obtain the following ML estimator:

$$\hat{\mathbf{\Omega}}_1 = \underset{\mathbf{\Omega}_1}{\operatorname{argmax}}\, \mathcal{L}(\mathbf{\Omega}_1), \quad (33)$$

where:

$$\mathcal{L}(\mathbf{\Omega}_1) = \frac{1 - |\nu_r|^2}{(1+|\nu_r|^2)r_{11}(\mathbf{\Omega}_1) - 2\Re e(\nu_r^* r_{12}(\mathbf{\Omega}_1))}. \quad (34)$$

As compared with the LS cost function that only depends on $r_{11}(\mathbf{\Omega}_1)$, the ML estimator also depends on the pseudo-variance of the noise $r_{12}(\mathbf{\Omega}_1)$.

B. Estimation of the transmitter IQ imbalance parameters

In this subsection, we show that the ML estimator of the transmitter (real-valued) IQ imbalance parameters $\mathbf{\Omega}_t$ can be obtained in closed form. The cost function $\mathcal{L}(\mathbf{\Omega}_1)$ is minimized with respect to the complex transmitter IQ parameters $\mathbf{\theta}_t$ if

$$\frac{\partial\mathcal{L}(\mathbf{\Omega}_1)}{\partial\mathbf{\theta}_t} \triangleq \begin{bmatrix} \frac{\partial\mathcal{L}(\mathbf{\Omega}_1)}{\partial\mu_t} \\ \frac{\partial\mathcal{L}(\mathbf{\Omega}_1)}{\partial\nu_t} \end{bmatrix} = \mathbf{0}, \quad (35)$$

where $\frac{\partial\mathcal{L}(\mathbf{\Omega}_1)}{\partial\alpha}$ corresponds to the complex derivative of a real-valued cost function [31]. Using the properties of the complex derivative, it can be checked that the derivative of the cost function is equal to $\mathbf{0}$ when:

$$(1+|\nu_r|^2)\frac{\partial r_{11}(\mathbf{\Omega}_1)}{\partial\mathbf{\theta}_t} - \nu_r^*\frac{\partial r_{12}(\mathbf{\Omega}_1)}{\partial\mathbf{\theta}_t} - \nu_r\frac{\partial r_{12}^*(\mathbf{\Omega}_1)}{\partial\mathbf{\theta}_t} = \mathbf{0} \quad (36)$$

where the derivatives of $r_{11}(\mathbf{\Omega})$ and $r_{12}(\mathbf{\Omega})$ are given by:

 Digital Object Identifier 10.32754/JMT.2022.2.04

$$\frac{\partial r_{11}(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t} = \frac{1}{N}\Big(\frac{\partial \mathbf{s}^T(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t}(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})^* + \frac{\partial \mathbf{s}^H(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t}(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})\Big), \quad (37)$$

$$\frac{\partial r_{12}(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t} = \frac{2}{N}\frac{\partial \mathbf{s}^T(\mathbf{\Omega}_1)}{d\boldsymbol{\theta}_t}(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y}), \quad (38)$$

$$\frac{\partial r_{12}^*(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t} = \frac{2}{N}\frac{\partial \mathbf{s}^H(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t}(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})^*. \quad (39)$$

The derivatives of $\mathbf{s}(\mathbf{\Omega}_1)$ can be expressed as:

$$\frac{\partial \mathbf{s}^T(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t} = \begin{bmatrix} \frac{\partial \mathbf{s}^T(\mathbf{\Omega}_1)}{\partial \mu_t} \\ \frac{\partial \mathbf{s}^T(\mathbf{\Omega}_1)}{\partial \nu_t} \end{bmatrix} = \mathbf{F}^T(\mathbf{\Omega}_h), \quad (40)$$

$$\frac{\partial \mathbf{s}^H(\mathbf{\Omega}_1)}{\partial \boldsymbol{\theta}_t} = \begin{bmatrix} \frac{\partial \mathbf{s}^H(\mathbf{\Omega}_1)}{\partial \mu_t} \\ \frac{\partial \mathbf{s}^H(\mathbf{\Omega}_1)}{\partial \nu_t} \end{bmatrix} = \nu_r^*\mathbf{F}^T(\mathbf{\Omega}_h). \quad (41)$$

Using these properties, it follows that the complex derivative is equal to zero if

$$\frac{1}{N}\mathbf{F}^T(\mathbf{\Omega}_h)((1+|\nu_r|^2)((\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})^* + \nu_r^*(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})) - -2\nu_r^*(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y}) - 2|\nu_r|^2(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})^*) = \mathbf{0}. \quad (42)$$

The term inside the parentheses can be factorized by $1-|\nu_r|^2$. Indeed, after some manipulations, the above equality can be arranged as:

$$\frac{1-|\nu_r|^2}{N}\mathbf{F}^T(\mathbf{\Omega}_h)((\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})^* - \nu_r^*(\mathbf{s}(\mathbf{\Omega}_1)-\mathbf{y})) = \mathbf{0}. \quad (43)$$

Using the assumption $|\nu_r| \neq 1$, this equality can be expressed under a simpler form as:

$$\mathbf{F}^H(\mathbf{\Omega}_h)(\mathbf{s}(\mathbf{\Omega}_1) - \nu_r\mathbf{s}^*(\mathbf{\Omega}_1)) = \mathbf{F}^H(\mathbf{\Omega}_h)(\mathbf{y}-\nu_r\mathbf{y}^*). \quad (44)$$

Then, as $\mathbf{s}(\mathbf{\Omega}_1) - \nu_r\mathbf{s}^*(\mathbf{\Omega}_1) = (1-|\nu_r|^2)\mathbf{F}(\mathbf{\Omega}_h)\boldsymbol{\theta}_t$, it follows that:

$$(1-|\nu_r|^2)\mathbf{F}^H(\mathbf{\Omega}_h)\mathbf{F}(\mathbf{\Omega}_h)\boldsymbol{\theta}_t = \mathbf{F}^H(\mathbf{\Omega}_h)(\mathbf{y}-\nu_r\mathbf{y}^*) \quad (45)$$

Finally, the estimator of the transmitter IQ imbalance parameters is simply given by:

$$\widehat{\boldsymbol{\theta}}_t = \frac{1}{1-|\nu_r|^2}\mathbf{F}^\dagger(\mathbf{\Omega}_h)(\mathbf{y}-\nu_r\mathbf{y}^*) \quad (46)$$

where $\mathbf{F}^\dagger(\mathbf{\Omega}_h) = (\mathbf{F}^H(\mathbf{\Omega}_h)\mathbf{F}(\mathbf{\Omega}_h))^{-1}\mathbf{F}^H(\mathbf{\Omega}_h)$ corresponds to the pseudo-inverse of $\mathbf{F}(\mathbf{\Omega}_h)$.

C. Estimation of the receiver IQ imbalance parameters and channel parameters

The estimation of $\mathbf{\Omega}_3 = [\mathbf{\Omega}_h^T, \mathbf{\Omega}_r^T]^T$ can be obtained by replacing the value of $\boldsymbol{\theta}_t$ by its ML estimate in (34). Let us introduce the orthogonal projector onto the kernel of $\mathbf{F}^H(\mathbf{\Omega}_h)$ as follows:

$$\mathbf{\Pi}_\mathbf{F}^\perp(\mathbf{\Omega}_h) \triangleq \mathbf{I} - \mathbf{F}(\mathbf{\Omega}_h)\mathbf{F}^\dagger(\mathbf{\Omega}_h). \quad (47)$$

By using the fact that $(1-|\nu_r|^2)\mathbf{y} = (\mathbf{y}-\nu_r\mathbf{y}^*) + \nu_r(\mathbf{y}-\nu_r\mathbf{y}^*)^*$ and the expression of the ML estimator of $\boldsymbol{\theta}_t$ in (46), the residual signal $\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_r, \mathbf{\Omega}_h, \widehat{\mathbf{\Omega}}_e)$ can be simplified as:

$$\begin{aligned}\mathbf{y} - \mathbf{s}(\mathbf{\Omega}_3, \widehat{\mathbf{\Omega}}_e) &= \mathbf{y} - \big(\mathbf{F}(\mathbf{\Omega}_h)\widehat{\boldsymbol{\theta}}_t + \nu_r\mathbf{F}^*(\mathbf{\Omega}_h)\widehat{\boldsymbol{\theta}}_t^*\big) = \\ &= \frac{1}{1-|\nu_r|^2}(\mathbf{\Pi}_\mathbf{F}^\perp(\mathbf{\Omega}_h)(\mathbf{y}-\nu_r\mathbf{y}^*) + \\ &+\nu_r\overset{*}{\mathbf{\Pi}_\mathbf{F}^\perp}(\mathbf{\Omega}_h)(\mathbf{y}-\nu_r\mathbf{y}^*)^*) = \\ &= \frac{1}{1-|\nu_r|^2}\mathbf{A}(\mathbf{\Omega}_h)\boldsymbol{\beta}(\nu_r). \end{aligned} \quad (48)$$

The matrix $\mathbf{A}(\mathbf{\Omega}_h)$ and the vector $\boldsymbol{\beta}(\nu_r)$ are defined as follows:

- $\mathbf{A}(\mathbf{\Omega}_h) \triangleq \big[\mathbf{\Pi}_\mathbf{F}^\perp(\mathbf{\Omega}_h)\mathbf{Y} \quad \overset{*}{\mathbf{\Pi}_\mathbf{F}^\perp}(\mathbf{\Omega}_h)\mathbf{Y}\big]$ is a $N \times 4$ matrix and $\mathbf{Y} = [\mathbf{y} \quad \mathbf{y}^*]$ is a $N \times 2$ matrix containing the received signal and its complex conjugate,
- $\boldsymbol{\beta}(\nu_r)$ is a 4×1 vector which is defined as:

$$\boldsymbol{\beta}(\nu_r) = \begin{bmatrix} 1 \\ -\nu_r \\ -|\nu_r|^2 \\ \nu_r \end{bmatrix}. \quad (49)$$

Note that the matrix $\mathbf{A}(\mathbf{\Omega}_h)$ satisfies the following property:

$$\mathbf{A}^*(\mathbf{\Omega}_h) = \mathbf{A}(\mathbf{\Omega}_h)\mathbf{P}, \quad (50)$$

where $\mathbf{P}$ is a permutation (and anti-diagonal) matrix defined by:

$$\mathbf{P} \triangleq \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \quad (51)$$

Using the definitions of $r_{11}(\mathbf{\Omega})$ and $r_{12}(\mathbf{\Omega})$ and the properties $\mathbf{A}^*(\mathbf{\Omega}_h) = \mathbf{A}(\mathbf{\Omega}_h)\mathbf{P}$ and $\mathbf{P}^T = \mathbf{P}$, the denominator of the cost function can be simplified as:

$$\begin{aligned}&(1+|\nu_r|^2)r_{11}(\mathbf{\Omega}_3, \widehat{\mathbf{\Omega}}_e) - 2\Re e\left(\nu_r^* r_{12}(\mathbf{\Omega}_3, \widehat{\mathbf{\Omega}}_e)\right) = \\ &= \frac{1}{N(1-|\nu_r|^2)^2}((1+|\nu_r|^2)\boldsymbol{\beta}^H(\nu_r)\mathbf{P}\mathbf{A}^T(\mathbf{\Omega}_h)\mathbf{A}(\mathbf{\Omega}_h)\boldsymbol{\beta}(\nu_r) - \\ &-\nu_r^*\boldsymbol{\beta}^T(\nu_r)\mathbf{A}^T\mathbf{A}\boldsymbol{\beta}(\nu_r) - \nu_r\boldsymbol{\beta}^H(\nu_r)\mathbf{P}\mathbf{A}^T(\mathbf{\Omega}_h)\mathbf{A}(\mathbf{\Omega}_h)\mathbf{P}\boldsymbol{\beta}^*(\nu_r)) = \\ &= \frac{1}{N(1-|\nu_r|^2)^2}((\boldsymbol{\beta}^H(\nu_r)\mathbf{P} - \nu_r^*\boldsymbol{\beta}^T(\nu_r))\mathbf{A}^T(\mathbf{\Omega}_h)\mathbf{A}(\mathbf{\Omega}_h)\boldsymbol{\beta}(\nu_r) - \\ &-\nu_r\boldsymbol{\beta}^H(\nu_r)\mathbf{P}\mathbf{A}^T(\mathbf{\Omega}_h)\mathbf{A}(\mathbf{\Omega}_h)(\mathbf{P}\boldsymbol{\beta}^*(\nu_r) - \nu_r^*\boldsymbol{\beta}(\nu_r))) = \\ &= \frac{1}{N(1-|\nu_r|^2)^2}\mathbf{v}^H(\nu_r)\mathbf{A}^T(\mathbf{\Omega}_h)\mathbf{A}(\mathbf{\Omega}_h)\mathbf{P}\mathbf{v}(\nu_r), \end{aligned} \quad (52)$$

where:

$$\mathbf{v}(\nu_r) \triangleq \mathbf{P}\boldsymbol{\beta}(\nu_r) - \nu_r\boldsymbol{\beta}^*(\nu_r). \quad (53)$$

Using the expression of $\boldsymbol{\beta}(\nu_r)$ and $\mathbf{P}$, it can be checked that $\mathbf{v}(\nu_r) = (1-|\nu_r|^2) \times [0,0,-\nu_r,1]^T$. Therefore, we obtain:

$$\begin{aligned}&(1+|\nu_r|^2)r_{11}(\mathbf{\Omega}_3, \widehat{\mathbf{\Omega}}_e) - 2\Re e\left(\nu_r^* r_{12}(\mathbf{\Omega}_3, \widehat{\mathbf{\Omega}}_e)\right) = \\ &= \frac{1}{N}[1 \quad -\nu_r^*]\mathbf{Y}^H\overset{H}{\mathbf{\Pi}_\mathbf{F}^\perp}(\mathbf{\Omega}_h)\mathbf{\Pi}_\mathbf{F}^\perp(\mathbf{\Omega}_h)\mathbf{Y}\begin{bmatrix} 1 \\ -\nu_r \end{bmatrix} = \\ &= (1+|\nu_r|^2)\mathbf{u}^H(\mathbf{\Omega}_r)\mathbf{R}(\mathbf{\Omega}_h)\mathbf{u}(\mathbf{\Omega}_r), \end{aligned} \quad (54)$$

where the 2×2 complex-valued matrix $\mathbf{R}(\boldsymbol{\Omega}_h)$ and the 2×1 complex-valued vector $\mathbf{u}(\boldsymbol{\Omega}_r)$ are defined as follows:

$$\mathbf{R}(\boldsymbol{\Omega}_h) \triangleq \frac{1}{N}\mathbf{Y}^H\boldsymbol{\Pi}_{\mathbf{F}}^{\perp H}(\boldsymbol{\Omega}_h)\boldsymbol{\Pi}_{\mathbf{F}}^{\perp}(\boldsymbol{\Omega}_h)\mathbf{Y}, \quad (55)$$

$$\mathbf{u}(\boldsymbol{\Omega}_r) \triangleq \frac{1}{\sqrt{1+|\nu_r|^2}}\begin{bmatrix}1\\-\nu_r\end{bmatrix}. \quad (56)$$

Note that the 2×1 vector $\mathbf{u}(\boldsymbol{\Omega}_r)$ has unit norm i.e., $\mathbf{u}^H(\boldsymbol{\Omega}_r)\mathbf{u}(\boldsymbol{\Omega}_r) = 1$. The ML estimation of $\boldsymbol{\Omega}_3$ can be obtained from the maximisation of $\mathcal{L}(\boldsymbol{\Omega}_3, \widehat{\boldsymbol{\Omega}}_e)$ or equivalently from the following optimization problem:

$$\{\widehat{\boldsymbol{\Omega}}_h, \widehat{\boldsymbol{\Omega}}_r\} = \arg\min_{\boldsymbol{\Omega}_h,\boldsymbol{\Omega}_r} \frac{1}{g(\boldsymbol{\Omega}_r)}\mathbf{u}^H(\boldsymbol{\Omega}_r)\mathbf{R}(\boldsymbol{\Omega}_h)\mathbf{u}(\boldsymbol{\Omega}_r), \quad (57)$$

where $g(\boldsymbol{\Omega}_r)$ corresponds to the receiver eccentricity factor and is defined as:

$$g(\boldsymbol{\Omega}_r) \triangleq \frac{1-|\nu_r|^2}{1+|\nu_r|^2}. \quad (58)$$

Therefore, similarly to the separable LS estimator, the ML estimator requires the minimization of a $(L+2)$-dimensional cost function. For small receiver IQ imbalance $g(\nu_r) \approx 1$, a low-complexity ML-based estimator can be obtained from the simpler constrained optimization problem:

$$\{\widehat{\boldsymbol{\Omega}}_h, \widehat{\boldsymbol{\Omega}}_r\} = \arg\min_{\boldsymbol{\Omega}_h,\boldsymbol{\Omega}_r} \mathbf{u}^H(\boldsymbol{\Omega}_r)\mathbf{R}(\boldsymbol{\Omega}_h)\mathbf{u}(\boldsymbol{\Omega}_r) \quad (59)$$

subject to

$$\mathbf{u}^H(\boldsymbol{\Omega}_r)\mathbf{u}(\boldsymbol{\Omega}_r) = 1. \quad (60)$$

Using this simple form, the estimation of $\boldsymbol{\Omega}_h$ and $\boldsymbol{\Omega}_r$ can be decoupled. Specifically, the estimation of the channel and receiver IQ parameters can be obtained from the smallest eigenvalue and eigenvector of $\mathbf{R}(\boldsymbol{\Omega}_h)$. Let us introduce the eigenvalue decomposition of the 2×2 matrix $\mathbf{R}(\boldsymbol{\Omega}_h)$ as follows:

$$\mathbf{R}(\boldsymbol{\Omega}_h) = \mathbf{U}(\boldsymbol{\Omega}_h)\boldsymbol{\lambda}(\boldsymbol{\Omega}_h)\mathbf{U}^H(\boldsymbol{\Omega}_h), \quad (61)$$

where:

- $\mathbf{U}(\boldsymbol{\Omega}_h) = [\mathbf{u}_1(\boldsymbol{\Omega}_h), \mathbf{u}_0(\boldsymbol{\Omega}_h)]$ is a 2×2 orthogonal matrix containing the 2 eigenvectors,
- $\boldsymbol{\lambda}(\boldsymbol{\Omega}_h) = \text{diag}(\lambda_1(\boldsymbol{\Omega}_h), \lambda_0(\boldsymbol{\Omega}_h))$ is a 2×2 diagonal matrix containing the eigenvalue value sorted in descending order ($\lambda_1(\boldsymbol{\Omega}_h) \geq \lambda_0(\boldsymbol{\Omega}_h)$).

In (59), the minimum of the cost function is attained when $\mathbf{u} = \mathbf{u}_0(\boldsymbol{\Omega}_h)$ and its minimal value is equal to $\lambda_0(\boldsymbol{\Omega}_h)$ [32]. Therefore, the channel parameters can be obtained from the following minimization problem:

$$\widehat{\boldsymbol{\Omega}}_h = \underset{\boldsymbol{\Omega}_h}{\text{argmin}}\, \lambda_0(\boldsymbol{\Omega}_h). \quad (62)$$

As $\mathbf{R}(\boldsymbol{\Omega}_h)$ is a 2×2 hermitian matrix, the smallest eigenvalue in (62) has a simple closed form expression given by (see for example [32]):

$$\lambda_0(\boldsymbol{\Omega}_h) = \frac{1}{2}\left(\text{tr}[\mathbf{R}(\boldsymbol{\Omega}_h)] - \sqrt{\text{tr}[\mathbf{R}(\boldsymbol{\Omega}_h)]^2 - 4\text{det}(\mathbf{R}(\boldsymbol{\Omega}_h))}\right) \quad (63)$$

where tr[.] corresponds to the trace of a matrix.

Note that when the channel parameters have been estimated, it is possible to extract the receiver IQ parameter ν_r from the eigenvector associated to the smallest eigenvalue of $\mathbf{R}(\widehat{\boldsymbol{\Omega}}_h)$, denoted $\widehat{\mathbf{u}}_0 = [\hat{u}_{00}, \hat{u}_{01}]^T$. Indeed, by imposing the equality $\widehat{\mathbf{u}}_0(\boldsymbol{\Omega}_h) = \widehat{\mathbf{u}}_0$, we obtain:

$$\hat{\nu}_r = -\frac{\hat{u}_{01}}{\hat{u}_{00}}. \quad (64)$$

The proposed estimation technique is finally summarized by the Algorithm 1. Similar to the LS algorithms, the most challenging step relies on the minimization of the multidimensional cost function. In the following, the Nelder-Mead simplex method is employed for the minimization of the ML cost function (62) [33]. Note that the development of a specific algorithm for the minimization of the cost function is out of the scope of this work.

V. Cramér-Rao Bounds

For any unbiased estimator, the Mean Squared Error is lower bounded by the Cramér-Rao Bounds (CRB) i.e. $\text{MSE}[\boldsymbol{\Omega}_k] = E[(\boldsymbol{\Omega}_k - \widehat{\boldsymbol{\Omega}}_k)^2] \geq \text{CRB}[\boldsymbol{\Omega}_k]$. In this section, we derive the CRB for the $L+7$ real-valued parameters $\boldsymbol{\Omega} = [\boldsymbol{\Omega}_h^T \;\; \boldsymbol{\Omega}_r^T \;\; \boldsymbol{\Omega}_t^T \;\; \sigma^2]^T$. In the literature, several expressions of the CRB have been derived for particular channel models under both transmitter and receiver IQ impairments [23], [26]. In this section, we consider the general setting where the channel can be described by the model in (2). Furthermore, as opposed to [23] and similarly to [26], we consider the more realistic setting where the noise is assumed circular before the receiver IQ impairments.

Algorithm 1 Maximum-Likelihood based Estimation of the channel parameters, transmitter and receiver IQ parameters.

Require: $\mathbf{y}$, $\mathbf{x}$,

1: Compute the two augmented matrices

$$\mathbf{X} = \begin{bmatrix}\mathbf{x} & \mathbf{x}^*\end{bmatrix}$$
$$\mathbf{Y} = \begin{bmatrix}\mathbf{y} & \mathbf{y}^*\end{bmatrix}$$

2: **Channel Estimation:** Find the value of $\widehat{\boldsymbol{\Omega}}_h$ that minimizes the smallest eigenvalue of

$$\mathbf{R}(\boldsymbol{\Omega}_h) = \frac{1}{N}\mathbf{Y}^H\boldsymbol{\Pi}_{\mathbf{F}}^{\perp H}(\boldsymbol{\Omega}_h)\boldsymbol{\Pi}_{\mathbf{F}}^{\perp}(\boldsymbol{\Omega}_h)\mathbf{Y}$$

where

$$\boldsymbol{\Pi}_{\mathbf{F}}^{\perp}(\boldsymbol{\Omega}_h) = \mathbf{I} - \mathbf{F}(\boldsymbol{\Omega}_h)\mathbf{F}^{\dagger}(\boldsymbol{\Omega}_h)$$

and

$$\mathbf{F}(\boldsymbol{\Omega}_h) = \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{X}.$$

3: **Receiver IQ parameter estimation:** Find $\widehat{\mathbf{u}}_0 = [\widehat{u}_{00}, \widehat{u}_{01}]^T$ the eigenvector associated to the smallest eigenvalue of $\mathbf{R}(\widehat{\boldsymbol{\Omega}}_h)$ and compute

$$\widehat{\nu}_r = -\frac{\widehat{u}_{01}}{\widehat{u}_{00}}$$

.

4: **Transmitter IQ parameter estimation:** Compute the complex-valued vector

$$\widehat{\boldsymbol{\theta}}_t = \begin{bmatrix}\widehat{\mu}_t\\ \widehat{\nu}_t\end{bmatrix} = \frac{1}{1-|\widehat{\nu}_r|^2}\mathbf{F}^{\dagger}(\widehat{\boldsymbol{\Omega}}_h)\left(\mathbf{y} - \widehat{\nu}_r\mathbf{y}^*\right)$$

5: **return** $\widehat{\boldsymbol{\Omega}}_h$, $\widehat{\boldsymbol{\Omega}}_r = [\Re e(\widehat{\nu}_r), \Im m(\widehat{\nu}_r)]^T$, and $\widehat{\boldsymbol{\Omega}}_t = [\Re e(\widehat{\boldsymbol{\theta}}_t^T), \Im m(\widehat{\boldsymbol{\theta}}_t^T)]^T$.

 Digital Object Identifier 10.32754/JMT.2022.2.04

Let us denote by $\boldsymbol{\Omega}_k$ the k^{th} element of the vector $\boldsymbol{\Omega}$. The Cramér-Rao Bounds (CRB) of $\boldsymbol{\Omega}_k$ is given by the k^{th} diagonal element of the inverse of the Fisher Information Matrix i.e.:

$$\text{CRB}[\boldsymbol{\Omega}_k] = [\mathbf{I}^{-1}(\boldsymbol{\Omega})]_{kk} \tag{65}$$

where $\mathbf{I}(\boldsymbol{\Omega})$ is the Fisher Information Matrix and $[.]_{kl}$ corresponds to the $(k,l)^{th}$ element of a matrix. As the augmented received vector is distributed as $\tilde{\mathbf{y}} \sim \mathcal{N}(\tilde{\mathbf{s}}(\boldsymbol{\Omega}_1), \mathbf{C}(\boldsymbol{\Omega}_2))$, the $(k,l)^{th}$ element of the Fisher Information Matrix is given by [29]:

$$\begin{aligned}[\mathbf{I}(\boldsymbol{\Omega})]_{kl} &= \left[\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}\right]^T \mathbf{C}^{-1}(\boldsymbol{\Omega}) \left[\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_l}\right] + \\ &+ \frac{1}{2}\text{tr}\left[\mathbf{C}^{-1}(\boldsymbol{\Omega})\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}\mathbf{C}^{-1}(\boldsymbol{\Omega})\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_l}\right]. \end{aligned} \tag{66}$$

Regarding the partial derivatives of $\tilde{\mathbf{s}}(\boldsymbol{\Omega})$, $\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial \sigma^2} = \mathbf{0}$ and:

$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_h]_k} = \begin{bmatrix}\Re e((1+\nu_r^*)\mathbf{G}_k\boldsymbol{\theta}_t) \\ \Im m((1-\nu_r^*)\mathbf{G}_k\boldsymbol{\theta}_t)\end{bmatrix}, \tag{67}$$

$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_r]_k} = \begin{bmatrix}\Re e(\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t) & \Im m(\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t) \\ -\Im m(\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t) & \Re e(\mathbf{F}(\boldsymbol{\Omega}_h)\boldsymbol{\theta}_t)\end{bmatrix}\mathbf{e}_k, \tag{68}$$

$$\frac{\partial \tilde{\mathbf{s}}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_t]_k} = \begin{bmatrix}\Re e((1+\nu_r^*)\mathbf{F}(\boldsymbol{\Omega}_h)) & -\Im m((1+\nu_r^*)\mathbf{F}(\boldsymbol{\Omega}_h)) \\ \Im m((1-\nu_r^*)\mathbf{F}(\boldsymbol{\Omega}_h)) & \Re e((1-\nu_r^*)\mathbf{F},(\boldsymbol{\Omega}_h))\end{bmatrix}\mathbf{e}_k, \tag{69}$$

where $\mathbf{e}_k$ is the unit column vector that contains only one 1 at the k^{th} row and zero elsewhere, and:

$$\mathbf{G}_k = \frac{\partial \mathbf{F}(\boldsymbol{\Omega}_h)}{\partial [\boldsymbol{\Omega}_h]_k} = \frac{\partial \mathbf{H}(\boldsymbol{\Omega}_h)}{\partial [\boldsymbol{\Omega}_h]_k}\mathbf{X}. \tag{70}$$

Concerning the covariance matrix $\mathbf{C}(\boldsymbol{\Omega})$, the derivatives $\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}]_k}$ are non-zero only for the receiver IQ parameters $\boldsymbol{\Omega}_r$ and for the noise variance σ^2. Mathematically, these non-zero derivatives are given by:

$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_r]_0} = \sigma^2 \begin{bmatrix}\Re e(\nu_r)+1 & 0 \\ 0 & \Re e(\nu_r)-1\end{bmatrix} \otimes \mathbf{I}_N, \tag{71}$$

$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial [\boldsymbol{\Omega}_r]_1} = \sigma^2 \begin{bmatrix}\Im m(\nu_r) & 1 \\ 1 & \Im m(\nu_r)\end{bmatrix} \otimes \mathbf{I}_N, \tag{72}$$

$$\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial \sigma^2} = \frac{1}{2}\mathbf{M}(\nu_r)\mathbf{M}^T(\nu_r) \otimes \mathbf{I}_N. \tag{73}$$

Note that when $\sigma^2 \to 0$, $\frac{\partial \mathbf{C}(\boldsymbol{\Omega})}{\partial \boldsymbol{\Omega}_r} = \mathbf{0}$.

VI. SIMULATION RESULTS

This section illustrates the performance of the proposed technique for several channel estimation problems under both transmitter and receiver IQ impairments. In this section, the proposed technique is compared with a simple LS estimator that assumes no IQ impairment, the separable LS estimator, and the LS estimator with constraint relaxation. These estimators are summarized in Table I. All these estimators have been implemented using Python NumPy / SciPy and their performances have been assessed using 5000 Monte Carlo simulations under different scenarios. In each simulation, the IQ imbalance parameters are arbitrary set to $\mu_t = 0.8+0.31j$, $\nu_t = 0.2-0.4j$, $\mu_r = 1$ and $\nu_r = 0.4+0.2j$. The deterministic transmitted samples $\mathbf{x}$ are randomly distributed over a QPSK constellation. Regarding the considered estimators, their cost functions are minimized using the Nelder-Mead algorithm. It should be emphasized that better algorithms can be employed for the minimization of the proposed cost function but, to be fair, we have preferred to compare the performance of the different estimator using the same classical optimization technique.

TABLE I. DESCRIPTION OF THE CONSIDERED ESTIMATORS

Label	Method	Cost Function	Nb dim.
simple	LS without IQ comp.	$\Vert \mathbf{y} - \mathbf{H}(\boldsymbol{\Omega}_h)\mathbf{x} \Vert^2$	L
LS	Separable LS	see (18)	L+2
LS relax	LS with Constraint Relax.	see (22)	L
Proposed	Approximate ML	see (62)	L

The column *Nb dim.* corresponds to the number of dimensions of the cost function.

A. *Estimation of the CFO*

In the first simulation, we have considered a L-taps FIR channel corrupted by a Carrier Frequency Offset (CFO). Mathematically, the channel model can therefore be described by:

$$\mathbf{H}(\boldsymbol{\Omega}_h) = \mathbf{D}(\omega)\mathbf{H}_0 \tag{74}$$

where the unknown channel parameter is $\boldsymbol{\Omega}_h = \omega$ and:

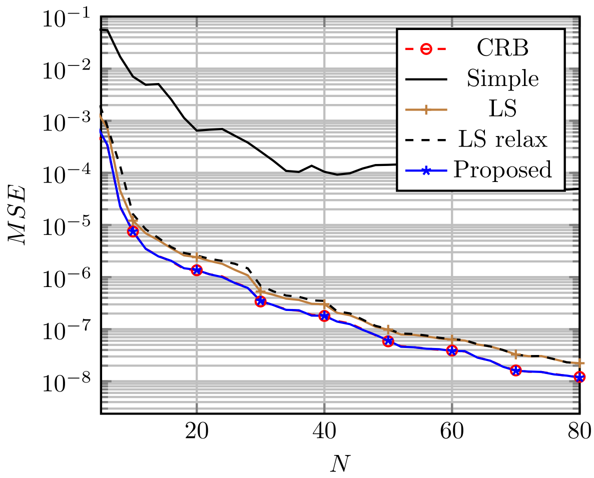
- $\mathbf{D}(\omega)$ is a $N \times N$ diagonal matrix with diagonal entries equal to $e^{j\omega n}$ $(n = 0,1,\cdots,N-1)$,
- $\mathbf{H}_0$ is a $N \times N$ Toeplitz matrix whose first column is equal to $\mathbf{h} = [h[0], h[1], \cdots, h[L]]^T$ (see for example (75)).

For this particular channel model, the CRB is computed using $\frac{\partial \mathbf{H}(\boldsymbol{\Omega}_h)}{\partial [\boldsymbol{\Omega}_h]_0} = \frac{\partial \mathbf{H}(\boldsymbol{\Omega}_h)}{\partial \omega} = \mathbf{D}_0\mathbf{H}(\boldsymbol{\Omega}_h)$ where $\mathbf{D}_0$ is $N \times N$ diagonal matrix defined as $\mathbf{D}_0 \triangleq \text{diag}(0, j, 2j, \cdots, (N-1)j)$. In each simulation, the non-zeros FIR coefficients are given by $h[0] = -2.32 - 0.875j$, $h[1] = 0.992 + 0.243j$, $h[2] = 0.182 - 0.516j$, $h[3] = 0.129 + 1.448j$. These coefficients are assumed perfectly known at the receiver side. Regarding the CFO, the frequency offset is arbitrary set to $\omega = 0.05$ radians/sample and the optimisation algorithms have been initialized with $\omega_{init} = 1.2 \times \omega$ radians/sample. Figure 2(a) presents the evolution of the Mean Squared Error, $E[(\omega - \widehat{\omega})^2]$, with respect to the number of samples N at $SNR = 20$dB. We observe that the proposed estimator is the only one that attains the CRB. Furthermore, we observe that the LS based estimators (LS and LS relax.) have similar performances for $N \geq 40$ samples and that the MSE obtained with these two estimators is roughly 1.5 times bigger than the one obtained with the proposed technique. We also see that the simple estimator is not able to provide a reliable estimate of the CFO. This clearly shows that the IQ parameters must be treated as nuisance parameters for channel estimation. Table II reports on the (median) computational time of each estimator. We observe that the computational time of the proposed technique is lower than the one required by the LS or LS relax techniques. Note that the high computational time obtained with the LS technique is mainly due to the fact that the minimization is performed along a higher-dimensional space.

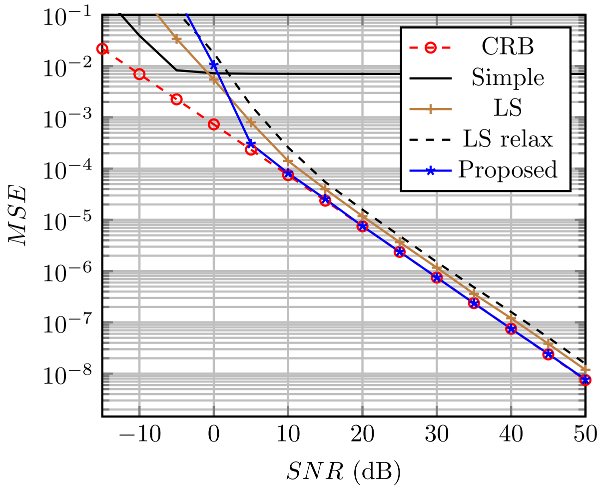
Digital Object Identifier 10.32754/JMT.2022.2.04

TABLE II. MEDIAN COMPUTATIONAL TIME VERSUS N IN SECONDS (PYTHON NUMPY/SCIPY CODE RUNNING ON A MAC MINITM 3.6 GHZ QUAD-CORE INTEL CORE I3 PROCESSOR)

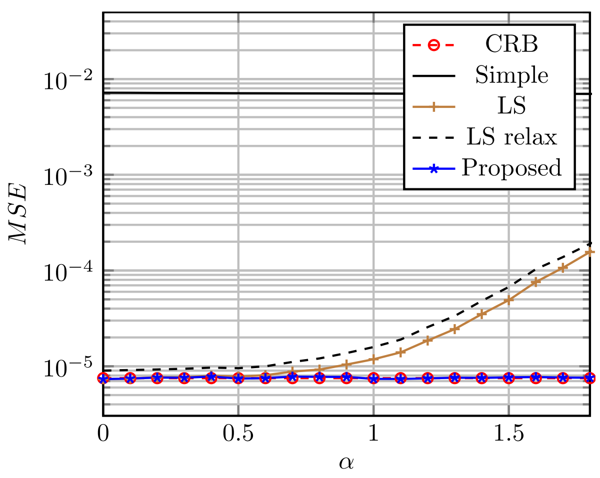
N	simple (s)	LS (s)	LS relax (s)	ML (s)
10	0.001821	0.020203	0.005563	0.005712
20	0.001876	0.037415	0.005150	0.005949
40	0.001776	0.122448	0.006359	0.007296
80	0.022560	0.182721	0.030564	0.010461



(a) MSE versus N (SNR = 20 dB)



(b) MSE versus SNR (N = 10 samples).



(c) MSE versus receiver impairment magnitude α (vr(α) = α×vr, N = 10 samples, SNR = 20 dB).

Figure 2. Estimation of the Carrier Frequency Offset (FIR parameters: [−2.32−0.875j, 0.992+0.243j, 0.182−0.516j, 0.129+1.448j])

Figure 2(b) shows the evolution of the MSE with respect to the SNR for a short preamble length ($N = 10$ samples). We see that the proposed technique attains the CRB for $SNR \geq 10$ dB. Furthermore, for $N = 10$ samples, we observe that the LS estimator outperforms the LS relax technique.

Finally, Figure 2(c) presents the evolution of the MSE with respect to the receiver impairment magnitude ν_r. In this simulation, the receiver imbalance is set to $\nu_r(\alpha) = \alpha\nu_r$, where the parameters $0 \leq \alpha \leq 1.5$ control the amount of imbalance. For $\nu_r = 0$, the performance of the LS and proposed estimator are roughly equivalent since the noise becomes circular[3]. Nevertheless, we observe that increasing the value of α leads to a significant performance degradation for the LS estimators, while the proposed technique is still optimal whatever the value of α.

B. Estimation of the FIR coefficients

In this subsection, we assume that the channel is described by a FIR channel with $L = 3$ coefficients. Under this assumption, the channel model can be expressed as:

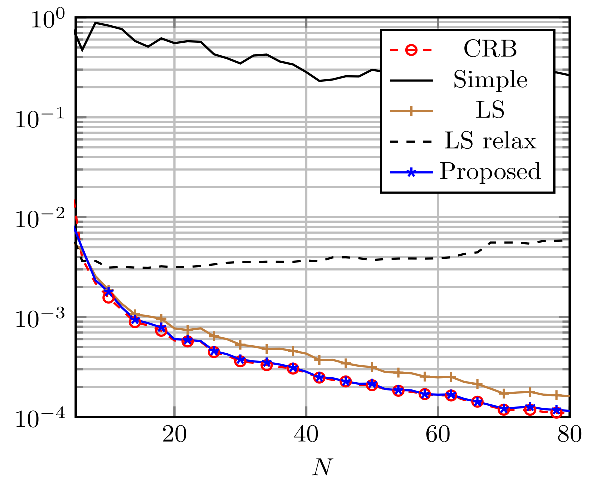
$$\mathbf{H}(\mathbf{\Omega}_h) = \begin{bmatrix} h[0] & 0 & \cdots & \cdots & \cdots & 0 \\ h[1] & h[0] & \ddots & & & \vdots \\ h[2] & h[1] & h[0] & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & h[2] & h[1] & h[0] \end{bmatrix} \quad (75)$$

When the FIR coefficients are unknown, it can be checked that the signal model have one additional complex indetermination. In this context, without lack of generality, we assume that the first FIR coefficient is set to $h[0] = 1$. The unknown real-valued parameters are then given by the 4×1 column vector $\mathbf{\Omega}_h = \tilde{\mathrm{h}} = [\Re e(\mathbf{h})^T, \Im m(\mathbf{h})]^T$, where $\mathbf{h} = [h[1], h[2]]^T$. For this particular channel model, the CRB is computed using $\frac{\partial \mathbf{H}(\mathbf{\Omega}_h)}{\partial [\mathbf{\Omega}_h]_0} = \mathbf{D}_0$, $\frac{\partial \mathbf{H}(\mathbf{\Omega}_h)}{\partial [\mathbf{\Omega}_h]_1} = \mathbf{D}_1$, $\frac{\partial \mathbf{H}(\mathbf{\Omega}_h)}{\partial [\mathbf{\Omega}_h]_2} = j\mathbf{D}_0$, and $\frac{\partial \mathbf{H}(\mathbf{\Omega}_h)}{\partial [\mathbf{\Omega}_h]_3} = j\mathbf{D}_1$ where $[\mathbf{D}_0]_{l,l-1} = 1$ and $[\mathbf{D}_1]_{l,l-2} = 1$. For $h[1]$ and $h[2]$, the CRB of the complex coefficients are respectively computed as $\mathrm{CRB}[h[1]] = \mathrm{CRB}[\mathbf{\Omega}_0] + \mathrm{CRB}[\mathbf{\Omega}_2]$ and $\mathrm{CRB}[h[2]] = \mathrm{CRB}[\mathbf{\Omega}_1] + \mathrm{CRB}[\mathbf{\Omega}_3]$. In the following simulations, the unknown FIR coefficients are arbitrary set to $h[1] = 0.992 + 0.243j$ and $h[2] = 0.182 - 0.516j$ and the optimisation algorithms have been initialized with $h_{init}[1] = 1.2 \times h[1]$ and $h_{init}[2] = 1.2 \times h[2]$.

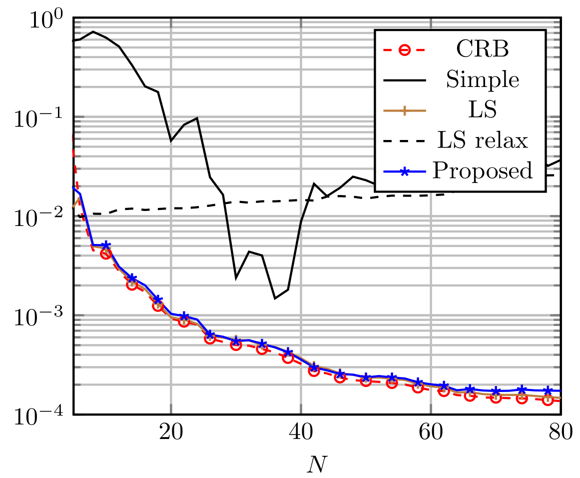
Figure 3 reports on the MSE of $h[1]$ and $h[2]$ versus N at SNR $= 25$ dB. For this setting, we observe that the simple and LS relax estimators do not provide a reliable estimate of the channel FIR coefficients, even for a large number of samples N. For $N \geq 50$ samples, the performance of the LS estimator of $h[2]$ is close to the one obtained with the proposed technique. For the FIR coefficient $h[1]$, we observe that our technique clearly outperforms the LS technique. For example,

[3]For circular Gaussian noise, the ML estimator reduces to the LS estimator.

 Digital Object Identifier 10.32754/JMT.2022.2.04

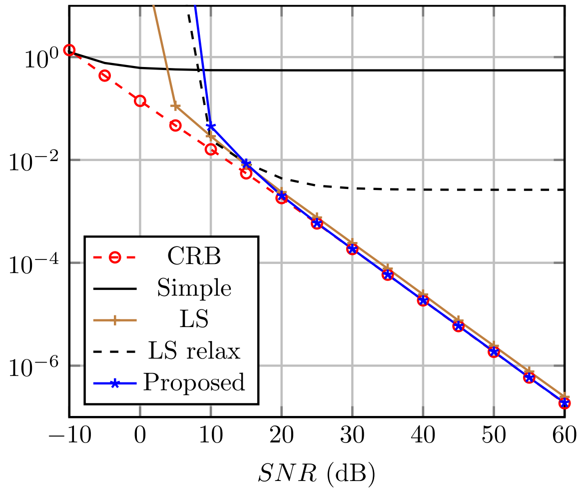


(a) Parameter h[1]

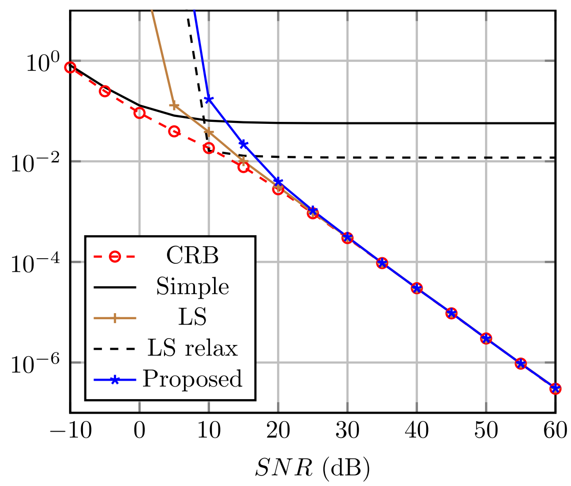


(b) Parameter h[2]

Figure 3. Estimation of the FIR parameters versus N (SNR = 25 dB, FIR parameters: [1,0.992 + 0.243j,0.182 − 0.516])



(a) Parameter h[1]



(b) Parameter h[1]

Figure 4. Estimation of the FIR parameters versus SNR (N = 20 samples, FIR parameters: [1,0.992 + 0.243j,0.182 – 0.516])

at $N = 80$ samples, the MSEs are respectively equal to $\text{MSE}[h_1] = 1.61 \times 10^{-4}$ and $\text{MSE}[h_1] = 1.14 \times 10^{-4}$ for the LS and proposed estimators. Note that, for this particular SNR, the samples between the CRB and the MSE of the proposed technique ($\text{MSE}[h_1] = 1.14 \times 10^{-4} > \text{CRB}[h_1] = 1.08 \times 10^{-4}$ and $\text{MSE}[h_2] = 1.73 \times 10^{-4} > \text{CRB}[h_2] = 1.36 \times 10^{-4}$).

Figure 4 presents the evolution of the MSE versus SNR for $N = 20$ samples. As previously reported, we note that the simple and LS relax estimators give poor performance. In particular, this figure clearly shows that these two estimators exhibit an error floor when the SNR is above 25 dB. Regarding the proposed estimator, we observe that the MSE attains the CRB for $SNR \geq 30$ dB. Furthermore, we see that the proposed technique outperforms the LS approach for the estimation of the FIR coefficient h [1], while the differences are more subtle for h [2].

VII. Conclusion

This paper proposed a new technique for the estimation of the channel parameters under both transmitter and receiver IQ impairments. The proposed estimator is based on the Maximum Likelihood technique and requires minimizing the eigenvalue of a 2×2 matrix with respect to the channel parameters only. As compared with the LS based techniques, simulation results highlighted the superiority of the proposed approach for the estimation of the CFO or channel FIR parameters in terms of statistical performance. More precisely, contrary to the LS estimators, simulations results showed that the proposed technique can attain the CRB asymptotically while having a significant lower computation complexity.

In future works, we will focus on the design of the transmitted signal in order to further reduce the computational complexity and/or to improve the statistical performance of the proposed estimator.

References

[1] M. Valkama, M. Renfors and V. Koivunen, “Advanced methods for I/Q imbalance compensation in communication receivers,” *IEEE Transactions on Signal Processing*, vol. 49, pp. 2335–2344, 2001.

[2] M. Valkama, M. Renfors and V. Koivunen, “Blind signal estimation in conjugate signal models with application to I/Q imbalance compensation,” *IEEE Signal Processing Letters*, vol. 12, pp. 733–736, 2005.

[3] L. Anttila, M. Valkama and M. Renfors, “Circularity-based I/Q imbalance compensation in wideband direct-conversion receivers,” *IEEE Transactions on Vehicular Technology*, vol. 57, pp. 2099–2113, 2008.

[4] D. Mattera, L. Paura and F. Sterle, "MMSE WL equalizer in presence of receiver IQ imbalance," *IEEE Transactions on Signal Processing*, vol. 56, pp. 1735–1740, 2008.

[5] J. Gao, X. Zhu, H. Lin and A. K. Nandi, "Independent component analysis based semi-blind I/Q imbalance compensation for MIMO OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 914–920, 2010.

[6] W. Nam, H. Roh, J. Lee and I. Kang, "Blind adaptive I/Q imbalance compensation algorithms for direct-conversion receivers," *IEEE Signal Processing Letters*, vol. 19, p. 475–478, 2012.

[7] J. Tubbax, B. Côme, L. Van der Perre, S. Donnay, M. Moonen and H. De Man, "Compensation of transmitter IQ imbalance for OFDM systems," in *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*, 2004.

[8] A. Tarighat, R. Bagheri and A. H. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3257–3268, 2005.

[9] A. Tarighat and A. H. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, pp. 240–247, 2007.

[10] K.-Y. Sung and C.-C. Chao, "Estimation and compensation of I/Q imbalance in OFDM direct-conversion receivers," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, pp. 438–453, 2009.

[11] A. Gomaa and L. M. A. Jalloul, "Data-aided I/Q imbalance estimation and compensation in OFDM systems," *IEEE Communications Letters*, vol. 18, pp. 459–462, 2014.

[12] N. Kolomvakis, M. Matthaiou, and M. Coldrey, "IQ imbalance in multiuser systems: Channel estimation and compensation," *IEEE Transactions on Communications*, vol. 64, pp. 3039–3051, 2016.

[13] J. Tubbax, B. Come, L. Van der Perre, S. Donnay, M. Engels, H. De Man and M. Moonen, "Compensation of IQ imbalance and phase noise in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 872–877, 2005.

[14] G. Xing, M. Shen and H. Liu, "Frequency offset and I/Q imbalance compensation for direct-conversion receivers," *IEEE Transactions on Wireless Communications*, vol. 4, p. 673–680, 2005.

[15] G.-T. Gil, I.-H. Sohn, J.-K. Park and Y. H. Lee, "Joint ML estimation of carrier frequency, channel, I/Q mismatch, and DC offset in communication receivers," *IEEE transactions on Vehicular Technology*, vol. 54, pp. 338–349, 2005.

[16] C.-J. Hsu, R. Cheng and W.-H. Sheen, "Joint least squares estimation of frequency, DC offset, IQ imbalance, and channel in MIMO receivers," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 2201–2213, 2008.

[17] F. Horlin, A. Bourdoux and L. Van der Perre, "Low-complexity EM-based joint acquisition of the carrier frequency offset and IQ imbalance," *IEEE Transactions on Wireless Communications*, vol. 7, p. 2212–2220, 2008.

[18] M. Inamori, A. M. Bostamam, Y. Sanada and H. Minami, "IQ imbalance compensation scheme in the presence of frequency offset and dynamic DC offset for a direct conversion receiver," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 2214–2220, 2009.

[19] J. Feigin and D. Brady, "Joint transmitter/receiver I/Q imbalance compensation for direct conversion OFDM in packet-switched multipath environments," *IEEE Transactions on Signal Processing*, vol. 57, pp. 4588–4593, 2009.

[20] B. Narasimhan, N. Al-Dhahir, M. Kocic, A. Yan, and Z. Zvonar, "Digital Compensation of Tx/Rx I/Q Imbalance in TD-SCDMA Systems," in *2011 IEEE Global Telecommunications Conference-GLOBECOM 2011*, 2011.

[21] M. Marey, M. Samir, and O. A. Dobre, "EM-based joint channel estimation and IQ imbalances for OFDM systems," *IEEE Transactions on Broadcasting*, vol. 58, pp. 106–113, 2012.

[22] W. Zhang, R. C. de Lamare, C. Pan, and M. Chen, "Joint TX/RX IQ imbalance parameter estimation using a generalized system model," in *2015 IEEE International Conference on Communications (ICC)*, 2015.

[23] H. Cheng, Y. Xia, Y. Huang, L. Yang and D. P. Mandic, "Joint Channel Estimation and Tx/Rx I/Q Imbalance Compensation for GFDM Systems," *IEEE Transactions on Wireless Communications*, vol. 18, pp. 1304–1317, 2019.

[24] D. Tandur and M. Moonen, "Joint adaptive compensation of transmitter and receiver IQ imbalance under carrier frequency offset in OFDM-based systems," *IEEE Transactions on Signal Processing*, vol. 55, pp. 5246–5252, 2007.

[25] C. Zhang, Z. Xiao, B. Gao, L. Su and D. Jin, "Three-Stage Treatment of TX/RX IQ Imbalance and Channel with CFO for SC-FDE Systems," *IEEE Communications Letters*, vol. 18, pp. 297-300, February 2014.

[26] X. Cai, Y.-C. Wu, H. Lin and K. Yamashita, "Estimation and compensation of CFO and I/Q imbalance in OFDM systems under timing ambiguity," *IEEE Transactions on Vehicular Technology*, vol. 60, pp. 1200–1205, 2011.

[27] C. Zhang, Z. Xiao, L. Su and D. Jin, "Joint estimation and compensation of TX and RX IQ imbalance for direct-conversion transceiver," *Wireless Personal Communications*, vol. 83, pp. 779–790, 2015.

[28] G. H. Golub and V. Pereyra, "The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate," *SIAM Journal on numerical analysis*, vol. 10, pp. 413–432, 1973.

[29] S. M. Kay, *Fundamentals of statistical signal processing*, Prentice Hall PTR, 1993.

[30] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Processing Letters*, vol. 15, pp. 841–844, 2008.

[31] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," in *IEE Proceedings H-Microwaves, Optics and Antennas*, 1983.

[32] R. A. Horn and C. R. Johnson, *Matrix analysis*, Cambridge University Press, 2012.

[33] F. Gao and L. Han, "Implementing the Nelder-Mead simplex algorithm with adaptive parameters," *Computational Optimization and Applications*, vol. 51, pp. 259–277, 2012.

 Digital Object Identifier 10.32754/JMT.2022.2.04