

Article

An Efficient Transmitter Feature Extraction Scheme with IQ Imbalance and Nonlinearity in TDD OFDM Systems

Yi Huang ^{1,2,*} , Aiqun Hu ¹, Jiayi Fan ², Huifeng Tian ², Xuebao Li ² and Yanfang Zheng ²¹ School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China; aqhu@seu.edu.cn² Zhangjiagang Campus, Jiangsu University of Science and Technology, Suzhou 215600, China; 999620150037@just.edu.cn (J.F.); thf830@just.edu.cn (H.T.); 2015000000482@just.edu.cn (X.L.); zyf068256@just.edu.cn (Y.Z.)

* Correspondence: hy2014@just.edu.cn

Abstract: Radio frequency (RF) fingerprints have been an emerging research topic for the last decade. Numerous algorithms for recognition have been proposed. However, very few algorithms for the accurate extraction of IQI and PA nonlinearity are available, especially when multiple paths are considered. In this study, we present a scheme that uses the transmitter in-phase/quadrature-phase imbalance (IQI) and the power amplifier (PA) nonlinearity as RF fingerprint features in time-division duplexing (TDD) OFDM systems, which are always considered to be harmful to data transmission. The scheme consists of two round trips with four steps for two cases: in the first, the IQI and PA nonlinearity are unknown at the terminal; in the second, they are known at the terminal. A channel state information (CSI)-tracking algorithm based on the sliding-window least squares method is first adopted at the terminal. In case A, the obtained CSI is sent to the base station (BS) to remove its impact there; in case B, this removal is conducted directly by using pre-equalization at the terminal. Then, by following a sequential iterative approach, the IQI and nonlinearity are individually calculated. Theoretical analyses reveal how CSI estimation errors influence subsequent estimates at the BS in these two cases. Furthermore, the approximate unbiasedness is verified. The theoretical variance and Cramer–Rao lower bound (CRLB) are also given. It is indicated that the theoretical minimum variance in case B is lower than that in case A from the perspective of the CRLB. The numerical results demonstrate the efficiency of the scheme in comparison with existing techniques in the literature.

Keywords: specific emitter identification; in-phase and quadrature-phase imbalance (IQI); nonlinearity; sliding-window least squares (LS); sequential iterative estimation; orthogonal frequency division multiplexing (OFDM)



Citation: Huang, Y.; Hu, A.; Fan, J.; Tian, H.; Li, X.; Zheng, Y. An Efficient Transmitter Feature Extraction Scheme with IQ Imbalance and Nonlinearity in TDD OFDM Systems. *Electronics* **2023**, *12*, 4108. <https://doi.org/10.3390/electronics12194108>

Academic Editor: Dimitra I. Kaklamani

Received: 26 July 2023

Revised: 8 September 2023

Accepted: 19 September 2023

Published: 30 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the development of 5G and 6G communication technologies, the physical-layer security of communication is becoming increasingly important, not only for future mobile communication systems but also for the internet of things (IoT) [1–3]. Security mechanisms that do not need to be built by using cryptography-based means in the upper layers and that rely solely on inherent physical-layer characteristics have their own unique characteristics, such as the inherent fingerprints of electronic components in transmitters, design defects or vulnerabilities in components in circuits, and the open propagation of electromagnetic waves in wireless communications [4]. In-phase/quadrature-phase imbalance (IQI) and power amplifier (PA) nonlinearity are two key transmitter characteristics that are relatively stable over the long term and vary from device to device [5,6]. They are harmful for data transmission, but are unique to each transmitter for use as a radio frequency (RF) fingerprint. Therefore, it is intuitive to use the transmitter IQI and PA nonlinearity as the RF fingerprint of a transmitter.

IQI estimation is always found in the field of channel equalization and is used to compensate for the distortion of the received signal generated by the IQI. A large number

of methods based on linear technology have already been extensively discussed. In [7,8], Alireza Tarighat proposed the least squares (LS) and adaptive equalization method for estimating the combination of transmitter and receiver IQI and channel state information (CSI) to compensate for the resulting transceiver performance degradation. The CSI and IQI were independently estimated based on the pilot by using linear estimation technology in a multiuser single-input–multiple-output (MU-SIMO) system; first, the linear minimum mean squared error (LMMSE) was used to estimate the CSI, and then LS was used to estimate the IQI, thus achieving the separation of the CSI and IQI parameters due to the assumption of IQI at the receiver only [9]. The transmitter (TX) and receiver (RX) IQI compensation problem in differential space–time block-coded orthogonal frequency division multiplexing (STBC-OFDM) systems was studied in [10] and developed further in [11]; the TX and RX IQIs were extracted by employing an algorithm based on a widely linear estimation. An iterative decision feedback (DFB) receiver was proposed to compensate for the IQI due to imperfections in the transceiver RF link for a single carrier with a frequency-domain equalization (SC-FDE) system [12]. This method was a nonlinear compensation method that could greatly improve the equalization performance of an SC-FDE system under IQI compared to the above-mentioned approaches. The authors of [13] used an iterative nonlinear least squares (NLS) method to jointly estimate the IQI in the TX, RX, and channels for generalized orthogonal frequency-division multiplexing (GFDM) systems. Ayse Elif Canbilen [14] investigated the impact of the IQI under generalized Beckmann fading channels for SM-MIMO systems, and both TX and RX IQIs were involved. A performance analysis of the effective channel estimation (e.g., the joint effect of the IQI and channel) and MMSE-based estimation was conducted in [15,16], where the RX IQI was considered in uplink MIMO-OFDM systems. Xiantao Cheng addressed two similar algorithms for estimating the IQI and CSI. One was a sequential estimation algorithm for determining the transmitter IQI and CSI [17], and the other was an EM-based estimation algorithm for determining the transceiver IQI and CSI [18]. Unlike previous methods, they were able to obtain separate estimates of the IQI and CSI. However, our simulation studies showed that these algorithms tended to fall into local minima.

The various methods mentioned above are able to effectively combat performance degradation due to IQI in direct conversion systems, but none of them take the PA nonlinearity of the transmitter into account. PA nonlinearity exists at the front-end of an RF link, and it distorts digitally modulated signals in wireless communications [19,20]. In addition to in-band distortion, the spectral regrowth of digitally modulated signals and adjacent channel interference are also produced by PA nonlinearity [19,20]. PA nonlinearity and IQI are always exploited in specific emitter identification (SEI) (also known as RF-distinct native attribute (RF-DNA) fingerprints) due to their uniqueness and long-term invariant nature [4].

In [21], Ming-Wei Liu extracted PA nonlinearity by using an alternate iterative LS estimation of the channel coefficients and distorted baseband signal to obtain this reliable characteristic in a transmitter based on a training sequence. However, this algorithm required a certain number of samples with a low-enough amplitude in the sequence; otherwise, severe performance degradation would be inevitable. Recently, many works have employed machine learning technology, such as classification algorithms and deep learning technology, to cope with SEI [5,6,22,23]. In such works, a number of simulated and practical trials were conducted, thus verifying the efficiency and reliability of their proposed methods. Nevertheless, unlike traditional approaches [7–18], machine learning is powerless in environments without samples, due to its requirements for learning.

In a word, in the previous literature, there are almost no algorithms that directly extract these two transmitter features separately, especially in the presence of multiple paths. The contributions of this study are summarized as follows:

- We establish a feature model of the IQI and PA nonlinearity at the transmitter with multiple paths. Following a brief analysis, we reveal that it is not possible to separately extract the parameters from just one transmission. However, separately extracting

IQI and PA nonlinearity under the impact of the CSI is necessary for robust feature extraction. To this end, we propose a scheme for separately estimating the IQI and PA nonlinearity, which involves two round trips with four steps.

- The CSI is estimated/tracked via a sliding-window LS algorithm based on previously depicted training sequences in the first round trip. Then, the IQI and PA nonlinearity of the transmitter are obtained by utilizing a sequential iterative approach after removing the impact of the CSI in the second round trip. The scheme includes two cases: A: IQI and PA nonlinearity are unknown at the terminal; B: IQI and PA nonlinearity are known at the terminal. These have different procedures in steps 3 and 4. The terminal needs to send the coded CSI to the base station (BS) in order to eliminate the channel effect at the BS for case A. The latter deals with the channel effect directly via pre-equalization in step 3 so as to estimate the IQI and PA nonlinearity more simply at the BS for case B.
- We also derive the performance expressions for the two cases and indicate the influence of CSI estimation/tracking. The results of the theoretical analysis show that the theoretical minimum variance in case B is lower than that in case A from the perspective of the Cramer–Rao lower bound (CRLB). In addition, the results of a numerical simulation verify the efficiency of the proposed scheme.

The rest of this study is organized as follows: We first describe the system model and give a CSI tracking algorithm based on the sliding-window LS (CTA_SWLS) with the IQI at the receiver of a terminal in Section 2. In Section 3, we propose an efficient scheme, with four steps, for eliminating the channel impact, and a sequential iterative approach for simultaneously and separately estimating the TX IQI and nonlinear coefficient at the BS is given. This is followed by a performance analysis and an analysis of the CRLB and complexity of the proposed scheme with the two cases in Section 4. Section 5 demonstrates the simulation results. Finally, conclusions are given in Section 6.

Notation: $(\bullet)^*$, $(\bullet)^T$, and $(\bullet)^H$ denote the complex conjugate, transpose, and complex conjugate transpose, respectively. \odot denotes element-by-element multiplication. $*$ represents circular convolution. $\text{DFT}\{x\}$ and $\text{IDFT}\{X\}$ denote the discrete Fourier transform of x and the inverse discrete Fourier transform of X , respectively. $X \bmod Y$ refers to taking the remainder after the division of X by Y . Imaginary units are represented by j . $\text{Re}\{X\}$ and $\text{Im}\{X\}$ denotes taking the real part and imaginary part of X , respectively. $\frac{X_{\bullet}}{Y_{\bullet}}$ indicates element-by-element division. The subscript or superscript i or k indicates the index of a variable, vector, or matrix in its context. $\frac{\partial \mathfrak{S}}{\partial x_0} = \frac{\partial \mathfrak{S}}{\partial x} \Big|_{x=x_0}$ refers to taking the partial derivative of \mathfrak{S} with respect to x when $x = x_0$.

2. System Model and Analysis

Here, we consider a direct upconversion transmitter in which the carrier frequency from an oscillation source is f_c , the phase difference between the I-branch and Q-branch is 90 degrees, and the amplitude coefficients are equal under ideal conditions. However, due to imperfections in the components of the analog circuits in the transmitter RF link and the diversity in manufacturing, the phases of the two branches do not have a difference of strictly 90 degrees, and the amplitude coefficients are not strictly the same, resulting in errors in the demodulated data on the receiver side if there is no countermeasure. On the other hand, the nonlinearity of the transmitter PA also leads to performance degradation at the receiver.

We assume here that no IQI and PA linearity are maintained across the operating range at the BS. Generally speaking, the BS has a more powerful function and can use higher-cost wide-linearity-region PA- and IQI-free devices; IQI and nonlinearity are even present in the RF link, which can also be pre-corrected to achieve more perfect IQ balance and PA linearity, while this is more difficult to achieve at the terminals due to the relatively low cost and performance of the devices, as well as the relatively large number of terminals.

2.1. TX IQI at the Terminal Only

Let $d(t) = d_I(t) + jd_Q(t)$ denote the ideal baseband equivalence symbol of the transmitter link at the terminal OFDM system. Because of the impact of the TX IQI at the terminal, this signal is distorted into the following form:

$$s(t) = ud(t) + vd^*(t), \tag{1}$$

where u and v are the TX IQI coefficients, which are defined by the following two equations:

$$\begin{aligned} u &= \cos(\theta/2) + j\alpha\sin(\theta/2) \\ v &= \alpha\cos(\theta/2) + j\sin(\theta/2), \end{aligned} \tag{2}$$

where α and θ are the amplitude imbalance and the phase imbalance, respectively, generated by the same local oscillator; this is called a symmetric IQI model. As shown in Figure 1a, under conditions of no IQI, $u = 1$ and $v = 0$.

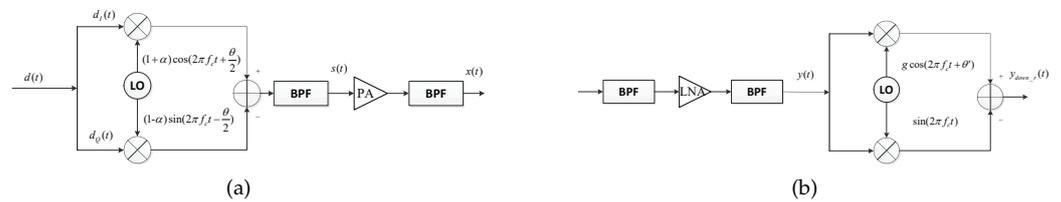


Figure 1. TX and RX block diagrams with IQI at the terminal. (a) Block diagram of the TX IQI at the terminal. (b) Block diagram of the RX IQI at the terminal.

Considering that the OFDM system consists of N subcarriers, the symbol period is T , $T = NT_s$, the cyclic prefix is N_{cp} , and T_s denotes the sampling period. The discrete signal expressions in the following parts of this study are all normalized by the sampling period unless specifically stated. The uplink signal received by the BS is

$$r(n) = [ud(n) + vd^*(n)] * h(n) + w(n), \tag{3}$$

where $h(n)$ is the multipath channel length of L , and $w(n)$ is additional white Gaussian noise (AWGN).

2.2. CSI Estimation with the RX IQI at the Terminal

If we consider the RX IQI at the terminal, a high-frequency (HF) signal that is received passes through a bandpass filter (BPF) and enters a low-noise amplifier (LNA), whose function is mainly to increase the output signal-to-noise ratio (SNR) (unlike the PA), so it can be considered linear over the entire operation range, as shown in Figure 1b. To facilitate the illustration of estimating the CSI at the terminal, we use an asymmetric model to describe the received signal in this section [9,24]. Thus, the signal received at the terminal in the time-division duplexing (TDD) system is

$$y_{down_r}(n) = \beta_1[d_B(n) * h(n) + w_{dwon_r}(n)] + \beta_2[d_B^*(n) * h^*(n) + w_{down_r}^*(n)], \tag{4}$$

where $d_B(n)$ represents a training sequence sent by the BS to the terminal, $w_{dwon_r}(n)$ is AWGN, and $w_{down_r}(n) \sim \mathcal{CN}(0, \sigma_{w_{down_r}}^2)$. $\beta_1 = \frac{1+ge^{j\theta}}{2}$ and $\beta_2 = \frac{1-ge^{-j\theta}}{2}$ are the RX IQI coefficients. Their complementary property is $\beta_1 + \beta_2^* = 1$, and this can be exploited to separate the received signal and the RX IQI. By taking the conjugate to (4) and adding it, we obtain

$$y_{down_r}(n) + y_{down_r}^*(n) = d_B^*(n) * h^*(n) + d_B(n) * h(n) + w_{down_r}(n) + w_{down_r}^*(n). \tag{5}$$

Let $\text{DFT}\{x(n)\} = X(k)$, which satisfies the conjugate symmetry $\text{DFT}\{x^*(n)\} = X^*(N - k)$. Taking the DFT of (5) and using conjugate symmetry and the complementary property results in

$$\begin{cases} Y_{\text{down}_r}(k) + Y_{\text{down}_r}^*(N - k) \\ = D_B(k)H(k) + D_B^*(N - k)H^*(N - k) \\ + W_{\text{down}_r}(k) + W_{\text{down}_r}^*(N - k) & 1 \leq k \leq \frac{N}{2} - 1 \\ \text{Re}\{Y_{\text{down}_r}(k)\} \\ = \text{Re}\{D_B(k)H(k) + W_{\text{down}_r}(k)\} & k = 0, \frac{N}{2}. \end{cases} \quad (6)$$

It can be seen in (6) that when sending $P (P \geq 2)$ different training sequences, we can obtain $\frac{N}{2} + 1$ independent linear equations; each equation can estimate two variables of the CSI via the LS. However, considering the time-varying nature of wireless channels, P sequences are required to be no longer than the channel coherence time [25].

$$P(T + T_{cp}) < T_c = \frac{0.423}{f_m}, \quad (7)$$

where f_m is the Doppler shift. For better tracking performance within the channel coherence time, we can send P consecutive but different training sequences circularly, herein utilizing the sliding-window LS algorithm to estimate the CSI for each symbol period. We introduce the symbol period index i and rewrite (6) in matrix form, resulting in

$$\mathbf{A}_k^i \mathbf{H}_k^i + \mathbf{W}_k^i = [\mathbf{A}_{k1}^i, \mathbf{A}_{k2}^i] \mathbf{H}_k^i + \mathbf{W}_k^i = \mathbf{Y}_k^i \quad P \leq i, 0 \leq k \leq \frac{N}{2}, \quad (8)$$

where

$$\mathbf{A}_{k1}^i = \begin{cases} [D_{B,(i-P+1) \bmod P}(k), D_{B,(i-P+2) \bmod P}(k), \dots, \\ D_{B,i \bmod P}(k)]^T & P \leq i, 1 \leq k \leq \frac{N}{2} - 1 \\ 2[\text{Re}\{D_{B,(i-P+1) \bmod P}(k)\}, \text{Re}\{D_{B,(i-P+2) \bmod P}(k)\}, \dots, \\ \text{Re}\{D_{B,i \bmod P}(k)\}]^T & P \leq i, k = 0, \frac{N}{2}, \end{cases} \quad (9)$$

$$\mathbf{A}_{k2}^i = \begin{cases} [D_{B,(i-P+1) \bmod P}^*(N - k), D_{B,(i-P+2) \bmod P}^*(N - k), \dots, \\ D_{B,i \bmod P}^*(N - k)]^T & P \leq i, 1 \leq k \leq \frac{N}{2} - 1 \\ 2[-\text{Im}\{D_{B,(i-P+1) \bmod P}(k)\}, -\text{Im}\{D_{B,(i-P+2) \bmod P}(k)\}, \dots, \\ -\text{Im}\{D_{B,i \bmod P}(k)\}]^T & P \leq i, k = 0, \frac{N}{2}, \end{cases} \quad (10)$$

$$\mathbf{H}_k^i = \begin{cases} [H^i(k), H^{i*}(N - k)]^T & P \leq i, 1 \leq k \leq \frac{N}{2} - 1 \\ [\text{Re}\{H^i(k)\}, \text{Im}\{H^i(k)\}]^T & P \leq i, k = 0, \frac{N}{2}, \end{cases} \quad (11)$$

$$\begin{aligned} \mathbf{Y}_k^i = & [Y_{\text{down}_r,(i-P+1)}(k) + Y_{\text{down}_r,(i-P+1)}^*(N - k), \\ & Y_{\text{down}_r,(i-P+2)}(k) + Y_{\text{down}_r,(i-P+2)}^*(N - k), \dots, \\ & Y_{\text{down}_r,i}(k) + Y_{\text{down}_r,i}^*(N - k)]^T \quad P \leq i, 0 \leq k \leq \frac{N}{2}, \end{aligned} \quad (12)$$

$$\begin{aligned} \mathbf{W}_k^i = & [W_{down_r,(i-P+1)}(k) + W_{down_r,(i-P+1)}^*(N-k), \\ & W_{down_r,(i-P+2)}(k) + W_{down_r,(i-P+2)}^*(N-k), \dots, \\ & W_{down_r,i}(k) + W_{down_r,i}^*(N-k)]^T P \leq i, 0 \leq k \leq \frac{N}{2}. \end{aligned} \tag{13}$$

It is easy to obtain the CSI estimate as follows:

$$\hat{\mathbf{H}}_k^i = (\mathbf{A}_k^{iH} \mathbf{A}_k^i)^{-1} \mathbf{A}_k^{iH} \mathbf{Y}_k^i \quad P \leq i, 0 \leq k \leq \frac{N}{2}. \tag{14}$$

Since interchanges of rows or columns of a matrix do not change its singular values, $(\mathbf{A}_k^{iH} \mathbf{A}_k^i)^{-1} = (\mathbf{A}_k^H \mathbf{A}_k)^{-1}$ is a fixed value. Let

$$\mathbf{Y}_k^i = \mathbf{A}_k^{iH} \mathbf{Y}_k^i = \mathbf{Y}_k^{i-1} + \mathbf{A}_k(imodP)^H [\mathbf{Y}_k^i(P-1) - \mathbf{Y}_k^{i-1}(0)] \quad P \leq i, \tag{15}$$

Thus,

$$\hat{\mathbf{H}}_k^i = \hat{\mathbf{H}}_k^{i-1} + (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \mathbf{Y}_k^i \quad P \leq i, 0 \leq k \leq \frac{N}{2}, \tag{16}$$

where $\mathbf{A}_k(i)$ denotes row $i + 1$ of \mathbf{A}_k , as well as $\mathbf{Y}_k^i(i)$. To avoid large estimation errors, it is necessary to design the training sequences with a small condition number for \mathbf{A}_k . The CSI tracking algorithm based on the sliding window LS (CTA_SWLS) is given in the form of Algorithm 1 and is shown in Figure 2.

Algorithm 1 CSI tracking algorithm based on the sliding-window LS (CTA_SWLS).

- 1: Initial $i = P - 1$;
 Calculate $\mathbf{B}_k = (\mathbf{A}_k^H \mathbf{A}_k)^{-1}$ and $\hat{\mathbf{H}}_k^i = \mathbf{B}_k \mathbf{A}_k^{iH} \mathbf{Y}_k^i$ offline.
 - 2: For $i \geq P$
 Calculate $\mathbf{Y}_k = \mathbf{A}_k(imodP)^H [\mathbf{Y}_k^i(P-1) - \mathbf{Y}_k^{i-1}(0)]$;
 Update $\hat{\mathbf{H}}_k^i = \hat{\mathbf{H}}_k^{i-1} + \mathbf{B}_k \mathbf{Y}_k$;
 $i = i + 1$
- End
-

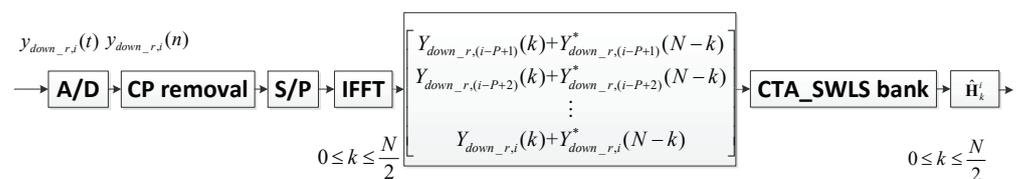


Figure 2. Block diagram of CSI tracking based on the sliding-window LS (CTA_SWLS) at the terminal.

2.3. TX Link with Nonlinear PA Only at the Terminal

Considering a memoryless PA, via Taylor expansions, a complex power-series expansion can be used to model the effect of a digital baseband signal passing through a transmitter front-end PA [20,21]. Accordingly, the even terms of the expanded series are discarded, leaving only the odd items, such that the baseband signal becomes

$$x = \sum_{i=0}^{M-1} b_m d^{m+1} (d^*)^m = \sum_{i=0}^{M-1} b_m |d|^{2m} d. \tag{17}$$

In (17), x denotes the signal after the baseband signal d passes the nonlinear PA, and b_m is a complex coefficient of odd items. Equation (17) shows that the items of $m = 0$ form

the linear component of the PA. To facilitate analysis, herein, let b_m be normalized to b_0 , so $b_0 = 1$. We rewrite (17) in matrix form as follows:

$$\mathbf{x} = [\mathbf{d}, \mathbf{D}_1]\mathbf{b} = \mathbf{D}\mathbf{b}, \tag{18}$$

where $\mathbf{x} = [x_0, x_1, \dots, x_{N-1}]^T$, $\mathbf{d} = [d_0, d_1, \dots, d_{N-1}]^T$, $[\mathbf{D}]_{i,m} = |d_i|^{2m}$ and $\mathbf{b} = [b_0, \dots, b_{M-1}]^T$.

2.4. Coexistence of the TX IQI and PA Nonlinearity in the Terminal Transmitter Link

By combining the IQ imbalance and PA nonlinearity and considering multipath channels and noise, the following received signals can be obtained at the BS:

$$\mathbf{r} = \mathbf{x} * \mathbf{h} + \mathbf{n} = ([\mathbf{s}, \mathbf{S}_R]\mathbf{b}) * \mathbf{h} + \mathbf{n} = (\mathbf{S}\mathbf{b}) * \mathbf{h} + \mathbf{n}, \tag{19}$$

where \mathbf{s} is the first column of \mathbf{S} , and \mathbf{S}_R is a matrix whose columns are composed of the last $M - 1$ columns of \mathbf{S} . \mathbf{n} is an additional white Gaussian noise (AWGN) vector, and $\mathbf{n} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_{N \times N})$. Thus, matrix \mathbf{S} becomes $[\mathbf{S}]_{i,m} = |ud_i + vd_i^*|^{2m} (ud_i + vd_i^*)$.

It can be seen from \mathbf{S} that the IQI coefficients u and v are in each element of \mathbf{S} , as well as in the nonlinearity coefficient vector \mathbf{b} ; all of these are the transmitter features that need to be extracted. If the CSI is unknown at the BS, an arbitrary set of values of u and v can be used to determine the matrix \mathbf{S} by combining arbitrary values of \mathbf{b} , provided that \mathbf{h} takes an appropriate value; this can make (19) consistent whether there is error or not. Hence, the parameters in this model cannot be determined if the CSI is unknown at the BS. With special cases in which $b_0 = 1, b_m = 0 (m > 0)$, which is a scenario that has been extensively studied in many previous papers [7–18]. However, for transmitter feature extraction, the requirement of obtaining independent estimates of the transmitter IQ imbalance and PA nonlinearity from multiple paths is necessary. Therefore, we need to seek a new approach to deal with this issue.

The situation will be different under the condition that the CSI is known according to (19). If a set of values of u and v are available, by using LS estimation, an estimate of \mathbf{b} can be obtained; the closer the values of u and v are to their true values in all estimates, the more accurate the estimate of \mathbf{b} is in the sense of the LS. Note that the LS algorithm requires that the column rank of \mathbf{s} be equal to the number of non-zero terms in \mathbf{b} ; otherwise, it will be ill conditioned, leading to large errors. Since \mathbf{S} can be written as

$$\mathbf{S} = \text{diag}(\mathbf{s})\Psi, \tag{20}$$

where $[\Psi]_{i,m} = |ud_i + vd_i^*|^{2m}$ is a Vandermonde matrix, $\text{rank}(\mathbf{S}) \leq \min\{\text{rank}(\text{diag}(\mathbf{s})), \text{rank}(\Psi)\}$ is equal to a non-zero number, and the rank of Ψ is equal to the number of different non-zero d_i values. In other words, the number of different non-zero d_i values in the training sequence is greater than the number of non-zero elements of \mathbf{b} . \mathbf{b} has an LS solution for the existence of IQI. Nevertheless, for IQI-free cases, the rank of Ψ is determined by the number of different non-zero $|d_i|$ values. It is, therefore, necessary that the number of different amplitudes is not less than M .

3. The Proposed Scheme

The above analysis shows that when the CSI is unknown at the BS, the model has no deterministic solution. Therefore, we propose a scheme for the joint and independent extraction of the IQI and nonlinear parameters in the terminal transmitter, and this scheme has two cases, which are shown in Figure 3.

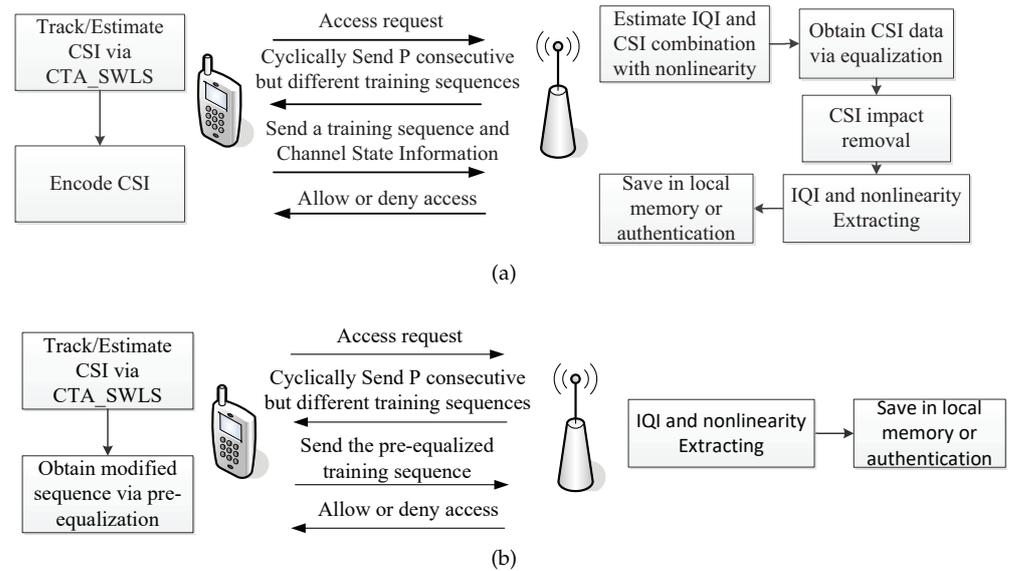


Figure 3. Block diagram of the proposed scheme. (a) Case A: IQI and PA nonlinearity are unknown at the terminal. (b) Case B: IQI and PA nonlinearity are known at the terminal.

3.1. Case A: IQI and PA Nonlinearity are Unknown at the Terminal

In this case, transmitter feature extraction and authentication are conducted via two round trips with four steps in TDD systems.

Step 1: The terminal sends an access request to the BS.

Step 2: Upon receipt of the request by the BS, P consecutive but different training sequences are cyclically sent to the terminal.

Step 3: The CTA_SWLS algorithm is employed to track the CSI after P training sequences are received at the terminal. When $\sum_{k=0}^{N/2} \|\hat{\mathbf{H}}_k^i - \hat{\mathbf{H}}_k^{i-1}\|^2 < \delta$ is satisfied (δ denotes a threshold), an FFT is performed on $\hat{\mathbf{H}}^i$ to obtain the time-domain CSI $\hat{\mathbf{h}}^i$. The first L ($L > \frac{\tau_{\max}}{T_s}$ is the max delay spread) values of $\hat{\mathbf{h}}^i$ are retained. Then, a training sequence is sent, and the CSI is truncated to the BS. It is assumed that the CSI remains constant during the last symbol period in step 2 and step 3. Note that to maintain the relevance of the CSI, the training sequence needs to be sent to the BS before the CSI.

Step 4: The BS estimates the combination of the CSI and transmitter IQI based on the training sequence in step 3 by using the LS or MMSE with PA nonlinearity; then, it demodulates the data, which are also known as the CSI. Followed by the removal of the effect of the CSI in the received signal, the BS then estimates the terminal IQI and PA nonlinearity coefficients. The features are stored in a local database either directly or after processing. It is also possible to compare data in the database by using pattern recognition so that the resulting signal to accept or reject the transmitter’s access can be given.

Considering the effects of PA nonlinearity and noise, in step 3, when the terminal encodes and modulates the estimated CSI, it should use a lower-order modulation method, such as BPSK or QPSK, on each OFDM subcarrier. For example, the number of effective paths is generally within 10 for a frequency-selective channel; when using 10 bits (1024-level uniform quantization) and BPSK modulation, a total of 100 bits are required. As long as the number of subcarriers satisfies $N \geq 100$, this requirement is met. After the BS receives the CSI and removes the impact of the channel, Equation (19) becomes

$$\mathbf{r}' = \mathbf{S}\mathbf{b} + \mathbf{n}', \tag{21}$$

where $\mathbf{r}' = \frac{1}{\sqrt{N}} \mathbf{F}^H \frac{\mathbf{F} \mathbf{r}}{\mathbf{F} \mathbf{h}_0}$ and $\mathbf{n}' = \frac{1}{\sqrt{N}} \mathbf{F}^H \frac{\mathbf{F} \mathbf{n}}{\mathbf{F} \mathbf{h}_0}$. $[\mathbf{F}]_{k,n} = \frac{1}{\sqrt{N}} e^{-j2\pi nk/N}$ is a Fourier matrix. This problem now turns into

$$\begin{aligned} & \min_{u,v,b} \|\mathbf{r}' - \mathbf{S}\mathbf{b}\|^2 \\ & s.t. \quad b_0 = 1. \end{aligned} \tag{22}$$

Let $\mathbf{v} = (\mathbf{b}, u, v)^T$. Since u and v in \mathbf{S} are to be estimated, \mathbf{b} is also to be estimated; we use a sequential iterative approach to estimate these quantities.

Setting the initial values to $u = 1$ and $v = 0$, first, we calculate

$$\hat{\mathbf{b}}' = (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \mathbf{S}_R^H (\mathbf{r}' - \mathbf{s}), \tag{23}$$

where $\hat{\mathbf{b}} = [1, \hat{\mathbf{b}}']^T$. Next, by utilizing the gradient descent method, the estimate of v can be iteratively obtained as follows:

$$\begin{aligned} \hat{v} &= \hat{v} - \mu \frac{\partial \|\mathbf{r}' - \mathbf{S}\hat{\mathbf{b}}\|^2}{\partial v} \\ &= \hat{v} - \mu \left(\frac{\partial (\hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{S} \hat{\mathbf{b}} - \mathbf{r}'^H \mathbf{S} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{r}')}{\partial v_x} \right. \\ &\quad \left. - j \frac{\partial (\hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{S} \hat{\mathbf{b}} - \mathbf{r}'^H \mathbf{S} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{r}')}{\partial v_y} \right), \end{aligned} \tag{24}$$

where μ is a small positive value for controlling the step size. In addition,

$$\begin{aligned} \frac{\partial (\hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{S} \hat{\mathbf{b}} - \mathbf{r}'^H \mathbf{S} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{r}')}{\partial v_x} &= \hat{\mathbf{b}}^H \left(\frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{S} + \mathbf{S}^H \frac{\partial \mathbf{S}}{\partial v_x} \right) \hat{\mathbf{b}} - \mathbf{r}'^H \frac{\partial \mathbf{S}}{\partial v_x} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{r}' \\ \frac{\partial (\hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{S} \hat{\mathbf{b}} - \mathbf{r}'^H \mathbf{S} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \mathbf{S}^H \mathbf{r}')}{\partial v_y} &= \hat{\mathbf{b}}^H \left(\frac{\partial \mathbf{S}^H}{\partial v_y} \mathbf{S} + \mathbf{S}^H \frac{\partial \mathbf{S}}{\partial v_y} \right) \hat{\mathbf{b}} - \mathbf{r}'^H \frac{\partial \mathbf{S}}{\partial v_y} \hat{\mathbf{b}} - \hat{\mathbf{b}}^H \frac{\partial \mathbf{S}^H}{\partial v_y} \mathbf{r}' \end{aligned}$$

where

$$\frac{\partial \mathbf{S}_{i,m}}{\partial v_x} = \left(\frac{\partial \mathbf{S}_{m,i}^H}{\partial v_x} \right)^* = \frac{\partial |ud_i + vd_i^*|^{2m} (ud_i + vd_i^*)}{\partial v_x} = |ud_i + vd_i^*|^{2m} \left[m \frac{2v_x |d_i|^2 + ud_i^2 + u^* d_i^{*2}}{(ud_i + vd_i^*)^*} + d_i^* \right], \tag{25}$$

$$\frac{\partial \mathbf{S}_{i,m}}{\partial v_y} = \left(\frac{\partial \mathbf{S}_{m,i}^H}{\partial v_y} \right)^* = \frac{\partial |ud_i + vd_i^*|^{2m} (ud_i + vd_i^*)}{\partial v_y} = |ud_i + vd_i^*|^{2m} \left[m \frac{2v_y |d_i|^2 - jud_i^2 + ju^* d_i^{*2}}{(ud_i + vd_i^*)^*} + jd_i^* \right]. \tag{26}$$

The following relationship of u and v based on (2) is used here to estimate u .

$$u = \sqrt{1 - v_y^2} + j \frac{v_x v_y}{\sqrt{1 - v_y^2}}. \tag{27}$$

After a finite number of iterations, the temporary suboptimal values of \hat{v} and \hat{u} can be obtained. Then, \hat{v} and \hat{u} are substituted into (23), and the iterative process is repeated until the stopping criteria are met. We summarize the estimation procedure in Algorithm 2.

Algorithm 2 IQI and nonlinearity sequential estimation algorithm (IQINSEA)

```

1: Initialization: Set  $u = 1, v = 0, i = 0, \mu, \varepsilon_1, \varepsilon_2$ 
2: For (external iteration)
    Update  $\hat{\mathbf{b}}'_i$  using (23);
    Form  $\hat{\mathbf{b}}_i = [1, \hat{\mathbf{b}}'_i]^T$ ;
    If  $\|[\hat{\mathbf{b}}_i, \hat{\vartheta}_i] - [\hat{\mathbf{b}}_{i-1}, \hat{\vartheta}_{i-1}]\| < \varepsilon_2$ 
        Stop and calculate  $\hat{u}_i$  using (27)
        Output:  $\hat{\mathbf{v}}_i$ .
    Else
         $k = 0$ ;
        For (internal iteration)
            Using (24) to update  $\hat{\vartheta}_{i,k+1}$ ;
            If  $\|\hat{\vartheta}_{i,k+1} - \hat{\vartheta}_{i,k}\| < \varepsilon_1$ 
                Stop
            Else
                 $k = k + 1$ ;
            Endif
        End
         $i = i + 1$ ;
    Endif
End

```

3.2. Case B: IQI and PA Nonlinearity are Known at the Terminal

Figure 3b demonstrates the procedure for case B. Here, prior knowledge is given at the terminal, and the corresponding processing is different to that in case A. The differences between case a and case B are as follows.

Step 3: The CTA_SWLS algorithm is employed to track the CSI after P training sequences are received at the terminal. When $\sum_{k=0}^{N/2} \|\hat{\mathbf{H}}_k^i - \hat{\mathbf{H}}_k^{i-1}\|^2 < \delta$ is satisfied, a modified sequence is directly obtained via a pre-equalization operation and sent to the BS instead of step 3 in case A.

Step 4: The BS receives and estimates the terminal IQI and nonlinear coefficients via Algorithm 2. The subsequent actions are the same as those in case A.

Since we only need to send a pre-equalization training sequence to the BS, case B is more suitable for faster-varying channels. With no CSI effect, the signal received by the BS should be shown as follows:

$$\mathbf{r}'' = \mathbf{S}\mathbf{b} + \mathbf{n}. \tag{28}$$

The matrix \mathbf{S} is known at both the terminal and the BS, but due to the CSI there is no unique solution, as described in the previous section. Therefore, we are able to send a pre-equalized training sequence to the BS after the CSI is estimated, and a signal with no CSI effect is equivalently received at the BS. To satisfy

$$\mathbf{r}'' = \mathbf{S}'\mathbf{b} * \mathbf{h} + \mathbf{n} = \mathbf{S}\mathbf{b} + \mathbf{n}, \tag{29}$$

that is,

$$\mathbf{S}'\mathbf{b} = \mathbf{F}^H \text{diag}(\mathbf{H})^{-1} \mathbf{F}\mathbf{S}\mathbf{b} = \tilde{\mathbf{r}}, \tag{30}$$

Equation (30) can be written as N independent equations for d_i :

$$\sum_{m=0}^{M-1} b_m |ud_i + vd_i^*|^{2m} (ud_i + vd_i^*) = \tilde{r}_i + \tilde{n}_i \quad 0 \leq i \leq N - 1 \tag{31}$$

or

$$F = \begin{cases} f_1 = \sum_{m=0}^{M-1} Q^m \Phi x_m - Cx = 0 \\ f_2 = \sum_{m=0}^{M-1} Q^m \Phi y_m - Cy = 0 \end{cases} \quad 0 \leq i \leq N - 1, \tag{32}$$

where

$$Q = |ud_i + vd_i^*|^2$$

$$\Phi x_m = bx_m(u_x d_{xi} - u_y d_{yi} + v_x d_{xi} + v_y d_{yi}) - by_m(u_x d_{yi} + u_y d_{xi} - v_x d_{yi} + v_y d_{xi})$$

$$\Phi y_m = by_m(u_x d_{xi} - u_y d_{yi} + v_x d_{xi} + v_y d_{yi}) + bx_m(u_x d_{yi} + u_y d_{xi} - v_x d_{yi} + v_y d_{xi})$$

$$Cx = \text{Re}\{\tilde{r}_i + \tilde{n}_i\}$$

$$Cy = \text{Im}\{\tilde{r}_i + \tilde{n}_i\}.$$

Equation (31) is a binary nonlinear system of equations with respect to d_{xi} and d_{yi} , and the solution can be obtained via the Newton–Raphson method. We omit the subscript i and let $\mathbf{d} = [d_x, d_y]^T$ so that we can obtain the following solution:

$$\mathbf{d}^{k+1} = \mathbf{d}^k - J(\mathbf{d}^k)^{-1}F(\mathbf{d}^k), \tag{33}$$

where $J(\mathbf{d}) = \begin{bmatrix} \frac{\partial f_1}{\partial d_x} & \frac{\partial f_1}{\partial d_y} \\ \frac{\partial f_2}{\partial d_x} & \frac{\partial f_2}{\partial d_y} \end{bmatrix}$ is a Jacobian matrix. In addition,

$$\frac{\partial f_1}{\partial d_x} = \sum_{m=0}^{M-1} \left(Q^m \frac{\partial \Phi x_m}{\partial d_x} + m Q^{m-1} \frac{\partial Q}{\partial d_x} \Phi x_m \right),$$

in which f_1 and Φx_m can be simultaneously replaced with f_2 and Φx_m . d_x can be replaced with d_y .

$$\begin{aligned} \frac{\partial Q}{\partial d_x} &= 2(u_x d_x - u_y d_y + v_x d_x + v_y d_y)(u_x + v_x) \\ &\quad + 2(u_x d_y + u_y d_x - v_x d_y + v_y d_x)(u_y + v_y) \end{aligned}$$

$$\begin{aligned} \frac{\partial Q}{\partial d_y} &= 2(u_x d_x - u_y d_y + v_x d_x + v_y d_y)(v_y - u_y) \\ &\quad + 2(u_x d_y + u_y d_x - v_x d_y + v_y d_x)(u_x - v_x) \end{aligned}$$

$$\frac{\partial \Phi x_m}{\partial d_x} = bx_m(u_x + v_x) - by_m(u_y + v_y)$$

$$\frac{\partial \Phi x_m}{\partial d_y} = bx_m(v_y - u_y) - by_m(u_x - v_x)$$

$$\frac{\partial \Phi y_m}{\partial d_x} = by_m(u_x + v_x) + bx_m(u_y + v_y)$$

$$\frac{\partial \Phi y_m}{\partial d_y} = by_m(v_y - u_y) + bx_m(u_x - v_x).$$

It is also possible to use nonlinear least squares to find the solution for the following optimization problem:

$$\min \left[\sum_{m=0}^{M-1} b_m |ud + vd^*|^2 (ud + vd^*) - (\tilde{r} + \tilde{n}) \right]^2. \tag{34}$$

It is given as

$$\mathbf{d}^{k+1} = \mathbf{d}^k - \eta \nabla(\mathbf{d}^k), \tag{35}$$

where $\nabla(\mathbf{d})=2J(\mathbf{d})^T F(\mathbf{d})$, and η is a small positive value for controlling the step size.

This is a system of N independent binary nonlinear equations that can be calculated in parallel at the transmitter side and implemented very quickly. The received signal that is transmitted to the BS can be eliminated from the CSI beforehand to facilitate subsequent processing at the BS. Once it is received by the BS, Algorithm 2 is used to directly estimate the IQI and nonlinear coefficients.

4. Performance Analysis

4.1. Performance Analysis of CTA_SWLS

Let \mathbf{H}_{0k}^i represent the true value of \mathbf{H}_k^i . Thus, the estimated CSI with k and $N - k$ subcarriers is

$$\hat{\mathbf{H}}_k^i = (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \sum_{p=0}^{P-1} \left\{ \mathbf{A}_k[(i-p) \bmod P]^H \left\{ \mathbf{A}_k[(i-p) \bmod P] \mathbf{H}_{0k}^{i-p} + \mathbf{W}_k(i-p) \right\} \right\}. \tag{36}$$

Taking the expectation to $\Delta \hat{\mathbf{H}}_k^i = \hat{\mathbf{H}}_k^i - \mathbf{H}_{0k}^i$ yields

$$\mathbb{E}[\Delta \hat{\mathbf{H}}_k^i] = (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \sum_{p=0}^{P-1} \left\{ \mathbf{A}_k[(i-p) \bmod P]^H \mathbf{A}_k[(i-p) \bmod P] (\mathbb{E}[\mathbf{H}_{0k}^{i-p}] - \mathbb{E}[\mathbf{H}_{0k}^i]) \right\}. \tag{37}$$

It can be seen from (37) that the unbiasedness of Algorithm 1 depends on the means of the past $P - 1$ true CSI. If the CSI is a stationary process over P symbol periods, then $\mathbb{E}[\Delta \hat{\mathbf{H}}_k^i]=0$, that is, Algorithm 1 is unbiased. It is assumed here that the channel is independent and identically distributed (IID), and $\mathbf{h}_0^i \sim \mathcal{CN}(\bar{\mathbf{h}}, \sigma_h^2 \mathbf{I}_{N \times N})$ for each path; thus, $\mathbf{H}_0^i = \text{DFT}(\mathbf{h}_0^i) \sim \mathcal{CN}(\bar{\mathbf{H}}, N\sigma_h^2 \mathbf{I}_{N \times N})$, where \mathbf{h}_0^i and \mathbf{H}_0^i denote the true CSI time-domain and frequency-domain vectors, respectively. $\bar{\mathbf{h}}$ and $\bar{\mathbf{H}}$ denote their respective means.

Then, the autocorrelation matrix of the channel error is given in the form of (38):

$$\begin{aligned} \mathbf{R}_k = \mathbb{E}[\Delta \hat{\mathbf{H}}_k^i \Delta \hat{\mathbf{H}}_k^{iH}] &= \underbrace{(\mathbf{A}_k^H \mathbf{A}_k)^{-1} \left\{ \sum_{p=0}^P \mathbf{A}_k(p)^H \mathbf{A}_k(p) \mathbf{A}_k(p)^H \mathbf{A}_k(p) \right\}}_{\text{I}} (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \sigma_k^2 + \sigma_k^2 \mathbf{I} \\ &- \underbrace{\left\{ (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \mathbf{A}_k[i \bmod P]^H \mathbf{A}_k[i \bmod P] + \mathbf{A}_k[i \bmod P]^H \mathbf{A}_k[i \bmod P] (\mathbf{A}_k^H \mathbf{A}_k)^{-1} \right\}}_{\text{II}} \sigma_k^2 + \underbrace{2(\mathbf{A}_k^H \mathbf{A}_k)^{-1} N \sigma_{w_{down,r}}^2}_{\text{III}}, \end{aligned} \tag{38}$$

where $\sigma_k^2 = \begin{cases} N\sigma_h^2 & 1 \leq k \leq \frac{N}{2} - 1 \\ N\sigma_h^2/2 & k = 0, \frac{N}{2} \end{cases}$. In (38), I is a fixed item, while II is an item with a period P . Both are generated by channel jitter, and III is generated by AWGN.

4.2. Error Propagation

In this subsection, the performance of Algorithm 2 in Step 4 is described. CSI estimation and quantization errors and demodulation errors in case A, as well as errors arising from pre-equalization in case B, are ignored. Before that, consideration of error propagation is necessary. In the scheme, error propagation affects the estimation of the terminal IQI and nonlinear parameters at the BS.

4.2.1. Case A

The received signal at the BS is

$$\mathbf{r} = \mathbf{S}_0 \mathbf{b}_0 * \mathbf{h}_0 + \mathbf{n}. \tag{39}$$

After removing the impact of the CSI, \mathbf{r} is now turned into $\hat{\mathbf{r}}'$ in the estimated version of (21). Then, an FFT is implemented on both sides:

$$\begin{aligned} \sqrt{N}\mathbf{F}\hat{\mathbf{r}}' &= \frac{\sqrt{N}\mathbf{F}\mathbf{r}_\bullet}{\sqrt{N}\mathbf{F}\hat{\mathbf{h}}} \\ &= \frac{\sqrt{N}\mathbf{F}\mathbf{S}_0\mathbf{b}_0}{\sqrt{N}\mathbf{F}\hat{\mathbf{h}}} \odot \sqrt{N}\mathbf{F}\mathbf{h}_0 + \frac{\sqrt{N}\mathbf{F}\mathbf{n}_\bullet}{\sqrt{N}\mathbf{F}\hat{\mathbf{h}}} \\ &= \left(\frac{\mathbf{1}^\circ}{\mathbf{1}^\circ + \mathbf{F}\Delta\mathbf{h}/\mathbf{F}\mathbf{h}_0}\right) \odot \sqrt{N}\mathbf{F}\mathbf{S}_0\mathbf{b}_0 \\ &\quad + \left(\frac{\mathbf{1}^\circ}{\mathbf{1}^\circ + \mathbf{F}\Delta\mathbf{h}/\mathbf{F}\mathbf{h}_0}\right) \odot \sqrt{N}\mathbf{F}\mathbf{n}_\bullet, \end{aligned} \tag{40}$$

where $\mathbf{1}_N^\circ = [1, 1, \dots, 1]_{1 \times N}^T$.

4.2.2. Case B

The pre-equalization effect satisfies

$$\hat{\mathbf{S}}'\mathbf{b}_0 * \hat{\mathbf{h}} + \mathbf{n} = (\mathbf{S}'_0 + \Delta\mathbf{S}')\mathbf{b}_0 * (\mathbf{h}_0 + \Delta\mathbf{h}) + \mathbf{n} = \mathbf{S}'_0\mathbf{b}_0 * \mathbf{h}_0 + \mathbf{n} = \mathbf{S}_0\mathbf{b}_0 + \mathbf{n} = \mathbf{r}'' \tag{41}$$

Then,

$$\hat{\mathbf{r}}'' = (\mathbf{S}'_0 + \Delta\mathbf{S}')\mathbf{b}_0 * \mathbf{h}_0 + \mathbf{n} = \mathbf{r}'' + \Delta\mathbf{S}'\mathbf{b}_0 * \mathbf{h}_0 = \mathbf{r}'' - \hat{\mathbf{S}}'\mathbf{b}_0 * \Delta\mathbf{h} \tag{42}$$

An FFT is implemented on both sides of the estimated version of (29):

$$\sqrt{N}\mathbf{F}\hat{\mathbf{r}}'' = \sqrt{N}\mathbf{F}\mathbf{r}'' - \sqrt{N}\mathbf{F}\hat{\mathbf{S}}'\mathbf{b}_0 \odot \sqrt{N}\mathbf{F}\Delta\mathbf{h} = \left(\frac{\mathbf{1}^\circ}{\mathbf{1}^\circ + \mathbf{F}\Delta\mathbf{h}/\mathbf{F}\mathbf{h}_0}\right) \odot \sqrt{N}\mathbf{F}\mathbf{S}_0\mathbf{b}_0 + \sqrt{N}\mathbf{F}\mathbf{n}_\bullet \tag{43}$$

It can be seen that apart from the noise items, the two received signals at the BS are equivalent to the convolution of the received signal and $\text{IDFT}\left(\frac{\mathbf{F}\mathbf{h}_0}{\mathbf{F}\mathbf{h}_0 + \mathbf{F}\Delta\mathbf{h}}\right)$ after removing the CSI in both case A and case B. The factor vector is determined by the CSI estimation error of CTA_SWLS and the true CSI value. Note that the pre-equalization does not introduce new errors in case B, it just propagates the estimation error of the CSI. The only difference between the two cases is the noise effect.

Each element of the factor vector $\frac{\mathbf{F}\mathbf{h}_0}{\mathbf{F}\mathbf{h}_0 + \mathbf{F}\Delta\mathbf{h}}$ follows a reciprocal normal distribution, with $\mathbf{F}\Delta\mathbf{h}$ obeying a gaussian distribution. Its expectation exists and is a scaled Faddeeva function whose exact expression depends on the sign of the imaginary part $\text{Im}(\mathbf{F}\mathbf{h}_0)$ [26]. Accordingly, the precise analysis of the factor vector's impact on the estimation error in Algorithm 2 is intractable. Here, we consider a moderate condition, $\Delta\mathbf{h} \approx 0$; then, $\frac{\mathbf{F}\mathbf{h}_0}{\mathbf{F}\mathbf{h}_0 + \mathbf{F}\Delta\mathbf{h}} \approx \mathbf{1}^\circ - \frac{\mathbf{F}\Delta\mathbf{h}}{\mathbf{F}\mathbf{h}_0}$.

4.3. Performance Analysis of Algorithm 2

A detailed mathematical analysis of the performance of Algorithm 2 in case A and case B can be found in Appendix A.

4.4. Complexity Analysis

Here, we consider a time-complexity analysis based on the number of multiplications. In Step 3, the FFTs of P training sequences and $\mathbf{B}_k = (\mathbf{A}_k^H \mathbf{A}_k)^{-1}$ are calculated offline. A total of P FFTs are required during the coherence time, and $\frac{N}{2} \log_2 N$ multiplications are acquired per FFT. There are $(\frac{N}{2} + 1)$ parallel calculations, with each requiring $[4P + \mathcal{O}(8)]$ multiplications.

However, during each symbol period, only \mathbf{Y}_k^i is needed for online calculation. Then, the iterations in Algorithm 1 are executed in parallel. Lastly, an IFFT is implemented for $\hat{\mathbf{h}}^i$. Therefore, a total of $N \log_2 N + 6$ multiplications are required online.

In case B, pre-equalization is a parallel calculation of a system of N equations. Each equation requires only $KM + 2$ complex multiplications, where K is the number of iterations of the Newton–Raphson method. For the implementation of Algorithm 2 in Step 4, $\{2K_1[(M^2 + 1)N + M(N + 2) + 4] + MN(2M + 1)\}K_2$ complex multiplications are needed. K_1 and K_2 denote the numbers of internal and external iterations in Algorithm 2, respectively.

5. Simulation and Analysis

In this section, simulations, analysis, and confirmation of the effectiveness of the proposed scheme are provided. The following simulation configuration was used: OFDM signals were sent at the BS and terminal, the number of subcarriers was $N = 1024$, $N_{cp} = 16$, $f_s = 1/T_s = 10$ MHz, and 64-QAM modulation was adopted in each subcarrier. The carrier frequency was $f_c = 2$ GHz, the relative movement velocity was 300 km/h, and $P = 7$ satisfied $P(T + T_{cp}) < T_c = \frac{0.423}{f_m} \approx 7.6 \times 10^{-4}$ s. In step 3 of case A, quantification errors were ignored, and the CSI was coded by using BPSK modulation for each subcarrier. Therefore, the BS demodulation error was not considered here. The time-varying channels were set up as independent Gaussian processes, and $\sigma_{w_{down,r}}^2 = \sigma^2$. The SNR was defined as follows:

$$\text{SNR} = \frac{\sum_{i=0}^{N-1} |d_i|^2}{N\sigma^2}. \quad (44)$$

The MSE was used as a performance indicator:

$$\text{MSE}[\hat{\mathbf{v}}] = 10 \log(\mathbb{E}[\|\hat{\mathbf{v}} - \mathbf{v}_0\|^2]). \quad (45)$$

The proposed scheme is labeled as 1 in subscript, the method from [17] is labeled as 2, and the method from [21] is labeled as 3. A total of 1000 independent trials were implemented per setting.

5.1. CSI Tracking via CTA_SWLS in Step 3

In this subsection, we simulate cyclically sending P consecutive but different training sequences to the terminal. $\beta_1(5^\circ, 0.3) = 0.5425 - j0.1438$, $\beta_2(5^\circ, 0.3) = 0.4575 - j0.1438$, $\bar{\mathbf{h}} = [1, 0.5, 0.2, 0.1, 0.3]^T$.

In Figure 4, one can see that the SNR difference did not significantly change the tracking performance. This performance was relatively more affected by the jitter of the channel itself. The reason for this was that the smoothing effect of RLS brought the tracking curve closer to the steady-state value of the channel, and the additive noise was effectively suppressed by using P training sequences to track the channel. Meanwhile, the time-varying nature of the P training sequences themselves had a larger cumulative impact on the tracking accuracy of the current symbol period. Thereby, the jitter of the channel had a greater impact on the tracking performance compared to the noise effect.

Figure 5 demonstrates the tracking performance of CTA_SWLS according to the MSE. Different noise levels only affected the mean value of the deviation, as shown by, for example, the fixed distance between the two curves at the top and bottom of Figure 5a,b, which was consistent with III in (38). However, the magnitude of the deviation was only affected by the channel jitter, which agreed with I and II in (38). In addition, the curves exhibited a cyclic pattern of the downlink with respect to the number of training sequences due to the cyclical sending of P sequences, which was in agreement with II of (38). Therefore, the theoretical variance of CTA_SWLS given by (38) was confirmed through this simulation.

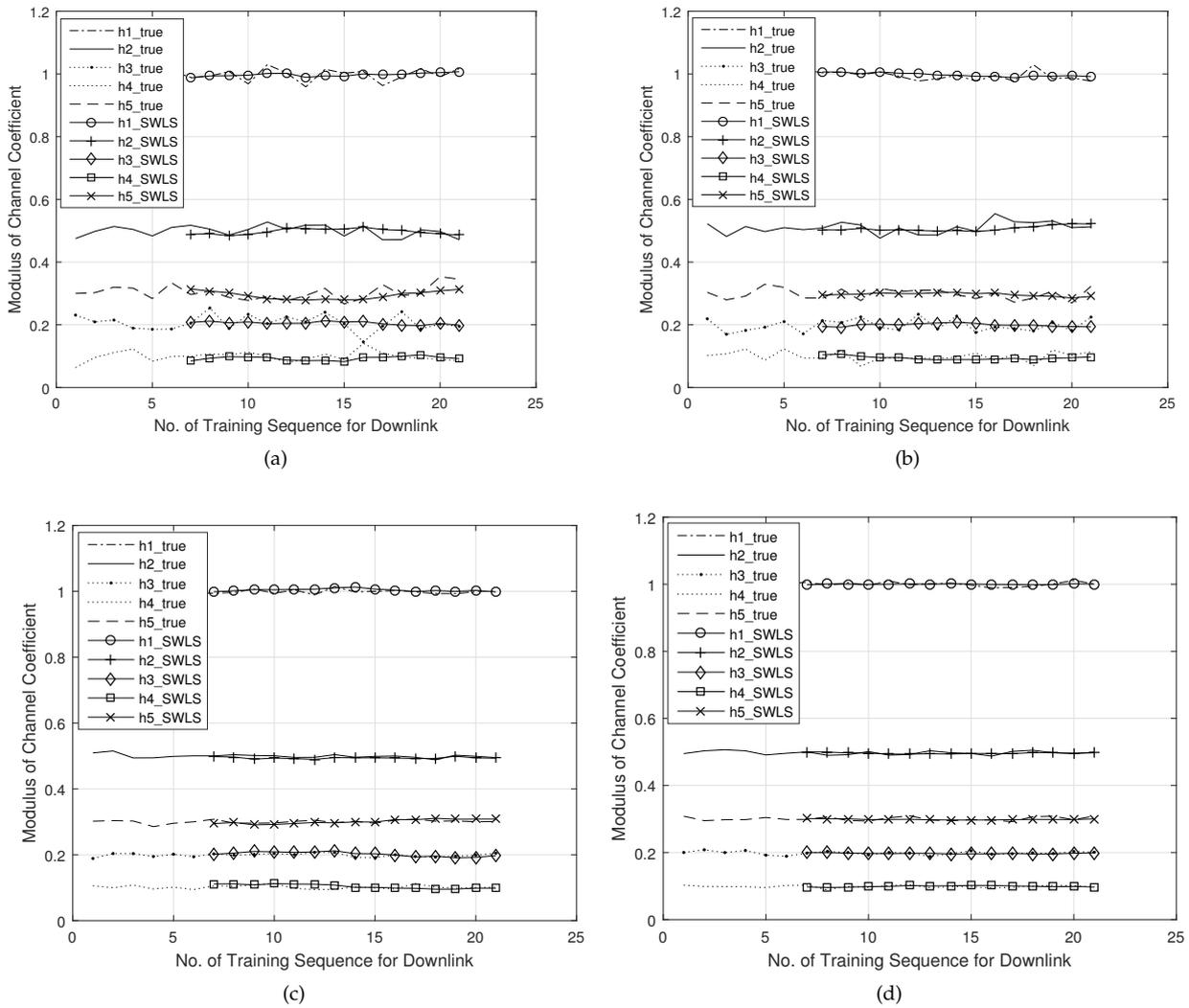


Figure 4. Tracking performance of CTA_SWLS for each channel tap. (a) $\sigma_h = 0.02$, SNR = 5 dB. (b) $\sigma_h = 0.02$, SNR = 20 dB. (c) $\sigma_h = 0.005$, SNR = 5 dB. (d) $\sigma_h = 0.005$, SNR = 20 dB.

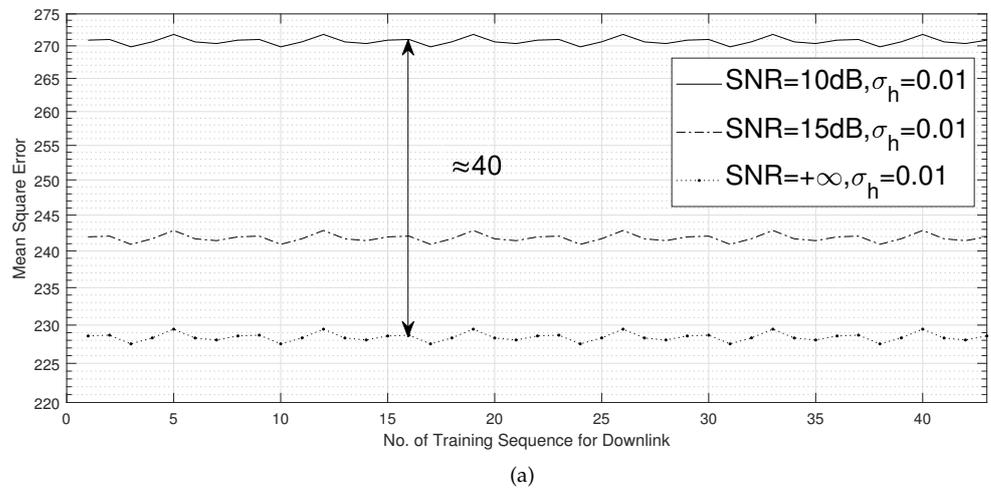


Figure 5. Cont.

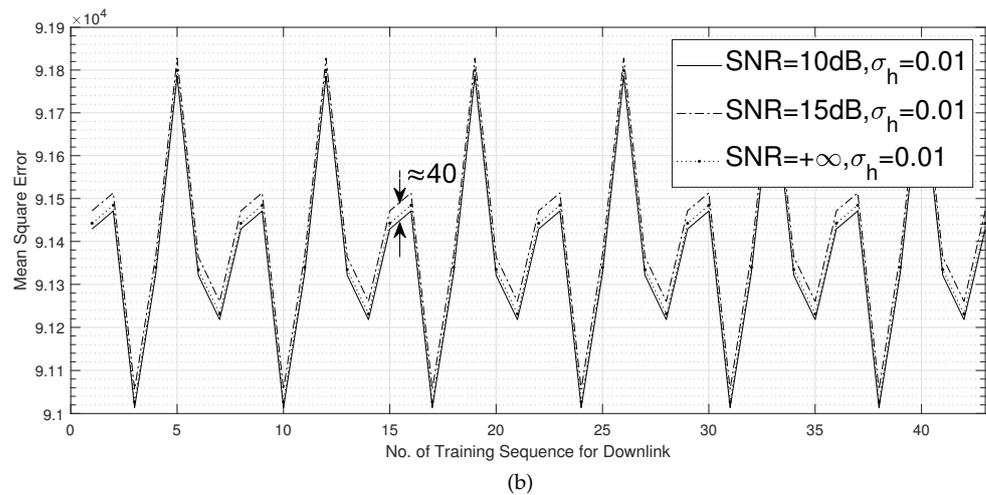


Figure 5. Tracking performance of CTA_SWLS according to the MSE. (a) $\sigma_h = 0.01$, SNR = 10 dB/15 dB/+ ∞ . (b) $\sigma_h = 0.2$, SNR = 10 dB/15 dB/+ ∞ .

5.2. IQI and PA Nonlinearity Estimation Performance via IQINSEA in Case A and Case B

Figures 6–8 show the MSE performance curves versus the SNR for IQINSEA in the two cases. The IQI and PA nonlinearity parameters were configured as follows: strong IQI and strong PA nonlinearity in Figure 6, strong IQI and weak PA nonlinearity in Figure 7, and weak IQI and strong PA nonlinearity in Figure 8. It was considered that the proposed scheme had to estimate the CSI first for both cases before the IQI and PA nonlinearity could be estimated by using IQINSEA. To facilitate fair comparisons, the ideal CSI was provided for method 2 and method 3; otherwise, the estimation performance of these two algorithms would be worse. Under this condition, since method 2 did not have the ability to estimate the PA nonlinearity and method 3 could not estimate the IQI, method 2 was only involved in the estimation of the IQI, and method 3 was only involved in the estimation of the PA nonlinearity.

In Figure 6, it can be observed that, in the strong IQI scenarios, the PA nonlinearity estimation with method 3 hardly improved as the SNR grew (Figures 6 and 7) because this method did not take the IQI into account, although there was some improvement in the weak IQI scenario (Figure 8). Accordingly, it can be seen that the effect of the IQI on the estimation performance of method 3 was relatively large. Note that the MSE performance of method 2 hardly varied with the SNR in all of the scenarios because this method is prone to local minima and does not consider the PA nonlinearity. These results were obtained with the assumption that the CSI was known at the BS for methods 2 and 3. Otherwise, the estimated errors would be even larger.

Nevertheless, of all of the scenarios, our proposed scheme was the most efficient among the three methods, i.e., it had a lower MSE and was closer to the CRLB. Another noteworthy aspect is that the MSE curves slowly decreased as the SNR grew, especially in the high-SNR regime. Note that the MSEs in case B were slightly lower than those in case A, especially with a low SNR, which verified the conclusion given in Section 4. This is because when the CSI was estimated, it was placed at the transmitter to directly remove the effect of the CSI on the system; then, the signal without the impact of the CSI was transmitted, i.e., case B. On the contrary, the CSI was further distorted in transmission, resulting in the removal of the impact of the CSI on the system at the receiver, thus making the error of the system larger, i.e., case A.

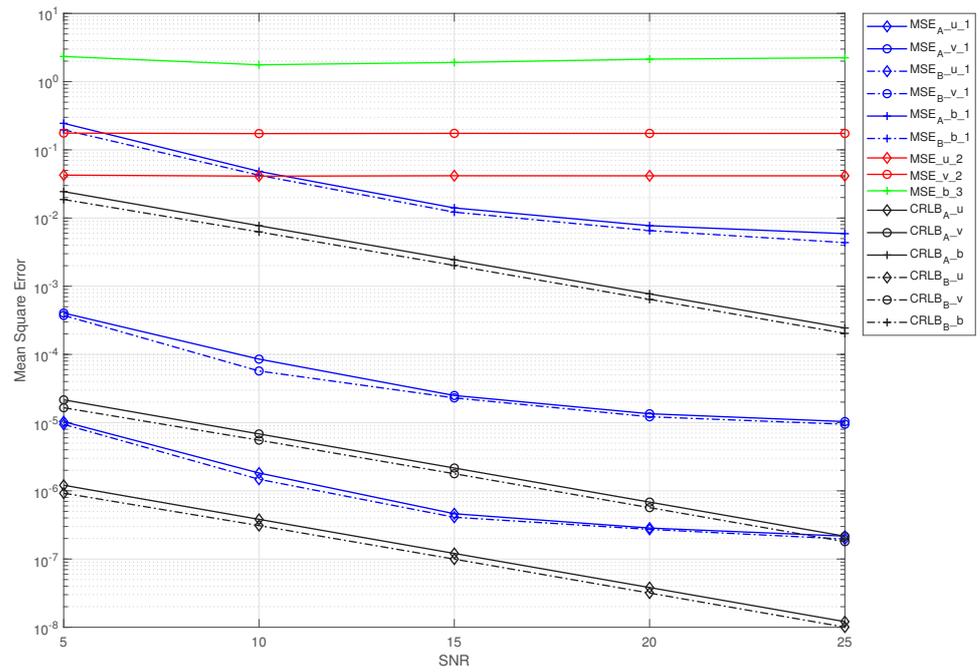


Figure 6. MSE curves for the three methods, as well as the CRLB. $\sigma_n = 0.02$. \mathbf{v} was set as follows: $\{[1, 0.3, 0.2 + j0.3, -j0.1], v(10^\circ, 0.2), u(10^\circ, 0.2)\}^T$.

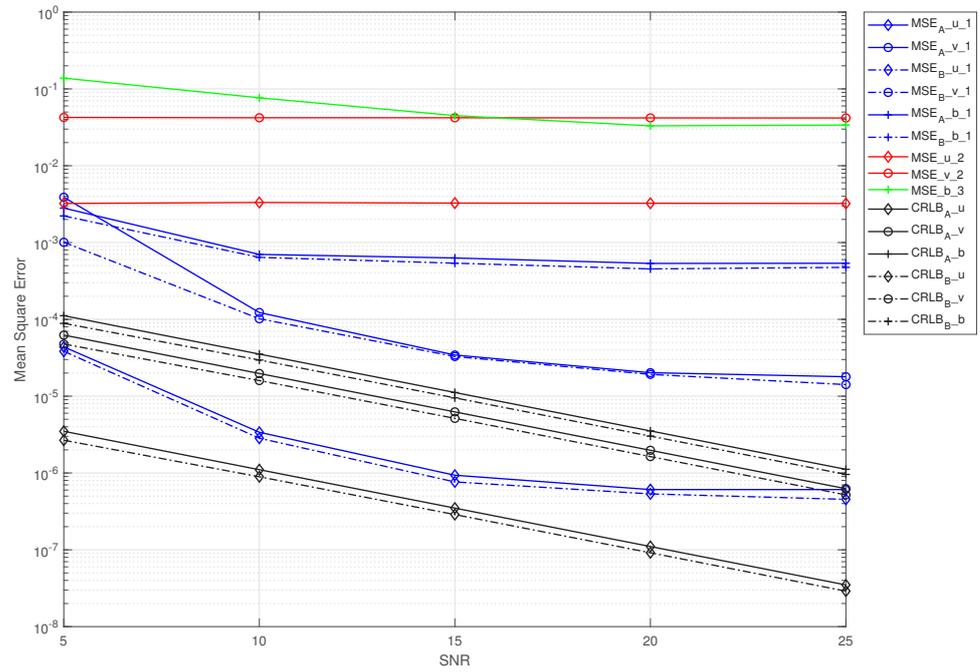


Figure 7. MSE curves for the three methods, as well as the CRLB. $\sigma_n = 0.02$. \mathbf{v} was set as follows: $\{[1, 0.4], v(10^\circ, 0.2), u(10^\circ, 0.2)\}^T$.

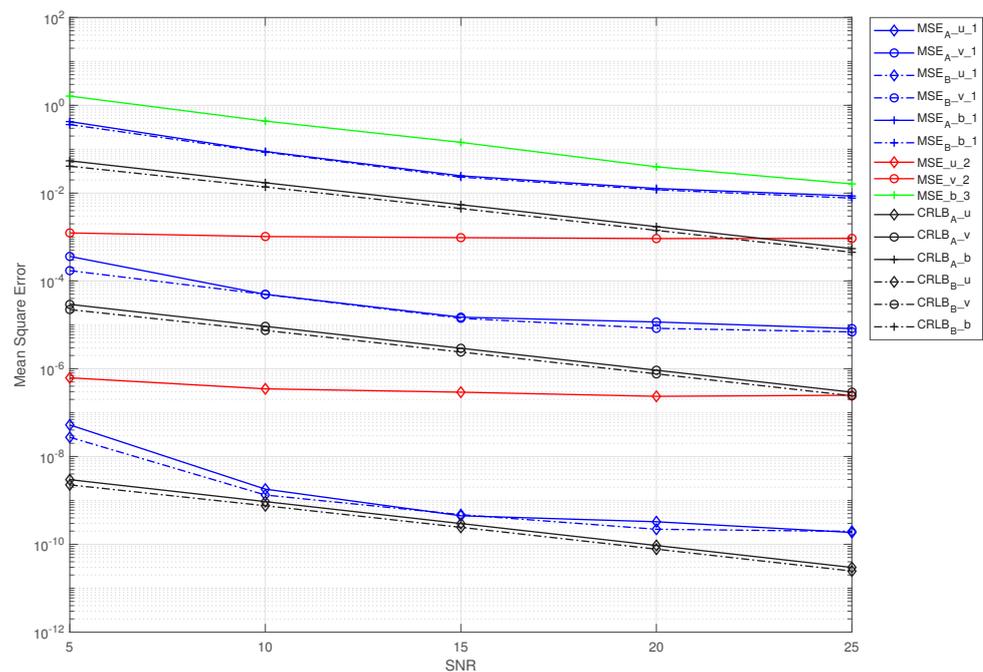


Figure 8. MSE curves for the three methods, as well as the CRLB. $\sigma_n = 0.02$. \mathbf{v} was set as follows: $\{[1, 0.3, 0.2 + j0.3, -j0.1], v(0.1^\circ, 0.01), u(0.1^\circ, 0.01)\}^T$.

6. Conclusions

In this study, two important features of a transmitter, i.e., IQI and PA nonlinearity, were utilized, and these are generated by the nonlinear circuit of the RF front-end at the terminal. They are detrimental to the efficiency of data transmission, but due to their uniqueness and long-term invariance, they can be used as RF fingerprints of a transmitter for terminal identification. This study focuses on the problem of extracting the IQI and PA nonlinearity of a transmitter efficiently and separately in the presence of multiple paths for TDD OFDM systems.

Via system modeling, it was found that only a single transmission could not yield a unique solution due to the coexistence of the IQI, PA nonlinearity, and multiple paths. To address this intractable problem, we proposed a scheme for independently estimating the IQI and nonlinearity, and this involved two round trips with four steps.

Once the training sequences from the BS were received, the real-time CSI was first estimated/tracked at the terminal by using the CTA_SWLS algorithm. The procedures were then divided into different parts for the two cases. In case A, the IQI and PA nonlinearity were unknown at the terminal; in case B, the IQI and PA nonlinearity were known at the terminal. In the two cases, the treatments for removing the impact of the CSI were distinct. Finally, the IQI and PA nonlinearity were estimated at the receiver by using IQINSEA.

With the theoretical analysis, the CSI tracking performance of CTA_SWLS was given. The influence of error propagation on further processing stages was given for both cases. Then, the approximate unbiasedness of the estimation algorithm was theoretically proven, and the variances of IQINSEA were given with the CRLB for both cases. The numerical results show the efficiency of the scheme in comparison with existing techniques in the literature.

Author Contributions: Conceptualization, Y.H. and A.H.; methodology, Y.H.; software, J.F.; validation, Y.H., X.L. and Y.Z.; formal analysis, Y.H. and H.T.; investigation, Y.H. and J.F.; resources, H.T.; data curation, Y.H. and H.T.; writing—original draft preparation, Y.H.; writing—review and editing, A.H.; visualization, X.L. and Y.Z.; supervision, A.H.; project administration, Y.H.; funding acquisition, A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by Jiangsu Provincial Key Laboratory of Network and Information Security (No. BM2003201), Natural Science Foundation of Jiangsu Province, China (Grant No. BK20201199)

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Performance Analysis of Algorithm 2

Appendix A.1. Case A

To facilitate the analysis of unbiasedness and efficiency, we analyze an estimator under the condition that the others are known. In practice, the parameters interact with each other, and the results of this analysis will have some errors with respect to its real performance. However, if the effects of interaction are considered, the analysis will be intractable due to its complexity. Considering the unbiasedness of v_x , it is first provided that v_y , u , and \mathbf{b} are known. Once the iterations converge to the optimal value, it is required that

$$\begin{aligned} & \frac{\partial(\mathbf{b}^H \mathbf{S}^H \mathbf{S} \mathbf{b} - \mathbf{r}'^H \mathbf{S} \mathbf{b} - \mathbf{b}^H \mathbf{S}^H \mathbf{r}')}{\partial v_x} \\ &= \mathbf{b}^H \left(\frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{S} + \mathbf{S}^H \frac{\partial \mathbf{S}}{\partial v_x} \right) \mathbf{b} - \mathbf{r}'^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{r}' \\ &= 0. \end{aligned} \tag{A1}$$

Substituting (21) into (A1), the following is given:

$$\begin{aligned} & \mathbf{b}^H \left\{ \frac{\partial \mathbf{S}^H}{\partial v_x} [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0] \right. \\ & \left. + [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0]^H \frac{\partial \mathbf{S}}{\partial v_x} \right\} \mathbf{b} \\ & - \mathbf{n}^H (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} \\ &= 0. \end{aligned} \tag{A2}$$

Since the presence of nonlinear components in the matrix \mathbf{S} ($b_m \neq 0, m > 1$) leads to the presence of higher-order quantities that cannot be easily and directly calculated, we use the Taylor approximation and retain the linear and quadratic terms.

$$[\mathbf{S} - \mathbf{S}_0]_{i,m} = \frac{\partial \mathbf{S}_{i,m}}{\partial v_{x0}} \Delta v_x + \frac{1}{2} \frac{\partial^2 \mathbf{S}_{i,m}}{\partial v_x^2} \Big|_{v_x = \tilde{\xi}_{i,m}} (\Delta v_x)^2, \tag{A3}$$

$$\left[\frac{\partial \mathbf{S}}{\partial v_x} \right]_{i,m} = \frac{\partial \mathbf{S}_{i,m}}{\partial v_{x0}} + \frac{\partial^2 \mathbf{S}_{i,m}}{\partial v_{x0}^2} \Delta v_x + \frac{\partial^3 \mathbf{S}_{i,m}}{\partial v_x^3} \Big|_{v_x = \tilde{\xi}_{i,m}} (\Delta v_x)^2, \tag{A4}$$

where $\tilde{\xi}_{i,m} \in (\min[v_x, v_{x0}], \max[v_x, v_{x0}])$. Let $\Delta \mathbf{H} = \sqrt{N} \mathbf{F} \Delta \mathbf{h}$, $\mathbf{H} = \sqrt{N} \mathbf{F} \mathbf{h}_0$, $\mathbf{\Lambda} = \text{diag}\{\frac{\Delta \mathbf{H}}{\mathbf{H}}\}$, and $\sigma_h^2 \approx 0$ to simplify the derivation, and (A2) becomes (A5).

$$\begin{aligned}
 & \mathbf{b}^H \left\{ \frac{\partial \mathbf{S}^H}{\partial v_x} [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0] + [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0]^H \frac{\partial \mathbf{S}}{\partial v_x} \right\} \mathbf{b} - \mathbf{n}^H (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} \\
 &= 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \Delta v_x + \mathbf{b}^H \left\{ \frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0 + (\mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0)^H \frac{\partial \mathbf{S}}{\partial v_x} \right\} \mathbf{b} - \mathbf{n}^H (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} + \mathcal{O}(\Delta v_x) \\
 &= 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \Delta v_x + \mathbf{b}^H \left\{ \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0 + (\mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0)^H \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \right\} \mathbf{b} \Delta v_x + \mathbf{b}^H \left\{ \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0 + (\mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0)^H \frac{\partial \mathbf{S}}{\partial v_{x0}} \right\} \mathbf{b} \\
 & - \mathbf{n}^H (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} - \mathbf{n}^H (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b} \Delta v_x - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} - \mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} \Delta v_x + \mathcal{O}(\Delta v_x) \\
 &= 0
 \end{aligned} \tag{A5}$$

Ignoring higher-order infinitesimals and taking the expectation on both sides of (A5), the following is obtained:

$$2\mathbf{b}^H \left(\frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \right) \mathbf{b} \mathbb{E}[\Delta v_x] \approx 0. \tag{A6}$$

This holds for any \mathbf{d} , and $\mathbb{E}[\Delta v_x]=0$ is obtained. So, under weak nonlinearity, the estimate of v_x is approximately unbiased. Similarly to the above derivation, v_y , u , and \mathbf{b} are also evidently unbiased. That is, Algorithm 2 is approximately unbiased in case A.

Let $\Xi = \mathbf{N} \mathbf{F} \mathbf{h} \mathbf{h}^H \mathbf{F}^H$, $\Xi' = \mathbf{N} \mathbf{F} \mathbf{h} \mathbf{h}^T \mathbf{F}^T$. Ignoring higher-order infinitesimals and taking the variance on both sides of (A5), $\mathbb{E}[\Delta v_x]=0$ and $\mathbb{E}[\mathbf{n}] = 0$ are substituted. \mathbf{v} and \mathbf{n} are independent. We obtain the variance of Δv_x through (A7).

$$\begin{aligned}
 \mathbb{D}_A[\Delta v_x] &= \mathbb{E}[\Delta v_x^2] \\
 &= \frac{2\mathbb{E}[\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b}] \sigma^2 + \mathbb{E}\{\mathbf{b}^H [\frac{\partial \mathbf{S}^H}{\partial v_{x0}} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0 + (\mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0)^H \frac{\partial \mathbf{S}}{\partial v_{x0}}] \mathbf{b}\}^2}{\left[2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \right]^2 + \mathbb{E}\{\mathbf{b}^H [\frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0 + (\mathbf{F}^H \mathbf{\Lambda} \mathbf{F} \mathbf{S}_0)^H \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2}] \mathbf{b}\}^2 + 2\mathbb{E}[\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b}] \sigma^2} \tag{A7} \\
 &\approx \frac{2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \sigma^2 + 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \sigma^2 + \Re_{v_x}}{\left[2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \right]^2 + 2\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b} \sigma^2 + 2\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b} \sigma^2 + \Re'_{v_x}}_{\mathbf{v}=\mathbf{v}_0}},
 \end{aligned}$$

where \Re_{v_x} and \Re'_{v_x} are given by

$$\begin{aligned}
 \Re_{v_x} &= 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \mathbf{F}^H (\mathbf{F} \mathbf{S}_0 \mathbf{b} \mathbf{b}^H \mathbf{S}_0^H \mathbf{F}^H \odot \mathbf{\Lambda}') \mathbf{F} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \\
 & + 2\text{Re}[\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \mathbf{F}^H (\mathbf{F} \mathbf{S}_0 \mathbf{b} \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \mathbf{F}^H \odot \mathbf{\Lambda}'') \mathbf{F} \mathbf{S}_0 \mathbf{b}],
 \end{aligned} \tag{A8}$$

$$\begin{aligned}
 \Re'_{v_x} &= 2\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \mathbf{F}^H (\mathbf{F} \mathbf{S}_0 \mathbf{b} \mathbf{b}^H \mathbf{S}_0^H \mathbf{F}^H \odot \mathbf{\Lambda}') \mathbf{F} \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b} \\
 & + 2\text{Re}[\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \mathbf{F}^H (\mathbf{F} \mathbf{S}_0 \mathbf{b} \mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \mathbf{F}^H \odot \mathbf{\Lambda}'') \mathbf{F} \mathbf{S}_0 \mathbf{b}],
 \end{aligned} \tag{A9}$$

where $\frac{\partial^2 \mathbf{S}_{i,m}}{\partial v_x^2}$ and $\frac{\partial^2 \mathbf{S}_{i,m}}{\partial v_y^2}$ are given in (A10) and (A11), respectively.

$$\begin{aligned}
 \frac{\partial^2 \mathbf{S}_{i,m}}{\partial v_x^2} &= m \left| \left(\sqrt{1 - v_y^2} + j \frac{v_x v_y}{\sqrt{1 - v_y^2}} \right) d_i + v d_i^* \right|^{2(m-1)} q_1 \left[\frac{m q_1}{(u d_i + v d_i^*)^*} + j \frac{v_y}{\sqrt{1 - v_y^2}} d_i + d_i^* \right] \\
 & + |u d_i + v d_i^*|^{2m} \left[\frac{m \frac{\partial q_1}{\partial v_x}}{(u d_i + v d_i^*)^*} - \frac{m q_1}{(u d_i + v d_i^*)^{*2}} \left(-j \frac{v_y}{\sqrt{1 - v_y^2}} d_i^* + d_i \right) \right] = \left(\frac{\partial^2 \mathbf{S}_{m,i}}{\partial v_x^2} \right)^*,
 \end{aligned} \tag{A10}$$

where $\mathbb{D}_A[\Delta v_y]$ is similar to $\mathbb{D}_A[\Delta v_x]$.

From (23) and (A2), it follows that

$$\hat{\mathbf{b}}'_A = (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \mathbf{S}_R^H [(\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0 \mathbf{b}_0 + (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{n} - \mathbf{s}]. \tag{A18}$$

The expectation is taken on both sides of

$$\mathbb{E}[\Delta \hat{\mathbf{b}}'_A] = \mathbb{E}[\hat{\mathbf{b}}'_A - \mathbf{b}'_0] = 0. \tag{A19}$$

Accordingly, the autocorrelation matrix of $\Delta \hat{\mathbf{b}}'_A$ is given as in (A20).

$$\begin{aligned} \mathbb{E}[\Delta \hat{\mathbf{b}}'_A \Delta \hat{\mathbf{b}}'^H_A] &= (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \mathbf{S}_R^H \mathbf{F}^H (\mathbf{F} \mathbf{S}_0 \mathbf{b}_0 \mathbf{b}_0^H \mathbf{S}_0^H \mathbf{F}^H \odot \mathbf{\Lambda}') \mathbf{F} \mathbf{S}_R (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \\ &+ (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \sigma^2 + (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \mathbf{S}_R^H \mathbf{F}^H \mathbf{\Lambda}' \mathbf{F} \mathbf{S}_R (\mathbf{S}_R^H \mathbf{S}_R)^{-1} \sigma^2 |_{\mathbf{v}=\mathbf{v}_0} \end{aligned} \tag{A20}$$

Note that, actually, the MSE will be larger than the above result if interactions are considered.

The variance of any unbiased estimator can only approximate the CRLB without restriction and not go below it; thus, we derive the CRLB of the model for comparative analyses here. The conditional probability density function (CPDF) of $\hat{\mathbf{r}}'$ is shown as in (A21).

$$\begin{aligned} \text{pdf}(\hat{\mathbf{r}}'; \mathbf{v}) &= \frac{1}{\pi^N \det[(\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})(\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H] \sigma^2} \exp \left\{ -\frac{1}{\sigma^2} [\hat{\mathbf{r}}' - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S} \mathbf{b}]^H [(\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})(\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F})^H]^{-1} \right. \\ &\quad \left. \times [\hat{\mathbf{r}}' - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S} \mathbf{b}] \right\} \end{aligned} \tag{A21}$$

Let $LF_A = \ln \text{pdf}(\hat{\mathbf{r}}'; \mathbf{v})$. It can be seen in (A21) that the CPDF satisfies the regularity conditions $\mathbb{E} \left[\frac{\partial LF_A}{\partial \mathbf{v}} \right] = 0$ among all of the available estimates. Accordingly, a minimum-variance unbiased estimator is attained, that is, the CRLB. The Fisher information matrix is obtained as

$$\mathbf{I}[\mathbf{v}]_A = -\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial v_i \partial v_j^*} \right]_{\mathbf{v}=\mathbf{v}_0}, \tag{A22}$$

where

$$-\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial v_x^2} \right] = \frac{2 \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b}}{\sigma^2} \tag{A23}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial v_y^2} \right] = \frac{2 \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b}}{\sigma^2} \tag{A24}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial \mathbf{b} \partial \mathbf{b}^H} \right] = \frac{\mathbf{S}_R^H \mathbf{S}_R}{\sigma^2} \tag{A25}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial v_x \partial v_y} \right] = \frac{\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b} + \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b}}{\sigma^2} \tag{A26}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial \mathbf{b}' \partial v_x} \right] = \frac{\mathbf{S}_R^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b}}{\sigma^2}, \quad -\mathbb{E} \left[\frac{\partial^2 LF_A}{\partial \mathbf{b}' \partial v_y} \right] = \frac{\mathbf{S}_R^H \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b}}{\sigma^2}. \tag{A27}$$

So, the CRLB in case A is equal to $\text{CRLB}_A = \mathbf{I}[\mathbf{v}]_A^{-1}$. Straightforwardly, the minimum variances of v and \mathbf{b}' in case A are given by

$$\text{CRLB}_A(v) = [\text{CRLB}_A]_{1,1} + [\text{CRLB}_A]_{2,2} \tag{A28}$$

$$\text{CRLB}_A[\mathbf{b}'] = [\text{CRLB}_A]_{3:M+1,3:M+1}.$$

Appendix A.2. Case B

Here, Equation (A2) becomes

$$\mathbf{b}^H \left\{ \frac{\partial \mathbf{S}^H}{\partial v_x} [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0] + [\mathbf{S} - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S}_0]^H \frac{\partial \mathbf{S}}{\partial v_x} \right\} \mathbf{b} - \mathbf{n}^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} - \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \mathbf{n} = 0. \tag{A29}$$

In a similar derivation to that for case A, we find that Algorithm 2 is also approximately unbiased in case B.

$$\mathbb{D}_B[\Delta v_x] = \mathbb{E}[\Delta v_x^2] = \frac{2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \sigma^2 + \Re_{v_x}}{\left[2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_{x0}} \frac{\partial \mathbf{S}}{\partial v_{x0}} \mathbf{b} \right]^2 + 2\mathbf{b}^H \frac{\partial^2 \mathbf{S}^H}{\partial v_{x0}^2} \frac{\partial^2 \mathbf{S}}{\partial v_{x0}^2} \mathbf{b} \sigma^2 + \Re'_{v_x}} \Bigg|_{\mathbf{v}=\mathbf{v}_0} \tag{A30}$$

$$\text{MSE}_B[\hat{\theta}] = \mathbb{D}_B[\Delta v_x] + \mathbb{D}_B[\Delta v_y], \tag{A31}$$

where $\mathbb{D}_B[\Delta v_y]$ is similar to $\mathbb{D}_B[\Delta v_x]$.

The CPDF of $\hat{\mathbf{r}}''$ is given as

$$\text{pdf}(\hat{\mathbf{r}}''; \mathbf{v}) = \frac{1}{\pi^N \sigma^2} \exp \left\{ -\frac{1}{\sigma^2} [\hat{\mathbf{r}}'' - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S} \mathbf{b}]^H [\hat{\mathbf{r}}'' - (\mathbf{I} - \mathbf{F}^H \mathbf{\Lambda} \mathbf{F}) \mathbf{S} \mathbf{b}] \right\}. \tag{A32}$$

Let $LF_B = \ln \text{pdf}(\hat{\mathbf{r}}''; \mathbf{v})$. After a similar derivation, this yields

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial v_x^2} \right] = \frac{2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} + 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b}}{\sigma^2} \tag{A33}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial v_y^2} \right] = \frac{2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b} + 2\mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b}}{\sigma^2} \tag{A34}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial \mathbf{b} \partial \mathbf{b}^H} \right] = \frac{\mathbf{S}_R^H \mathbf{S}_R + \mathbf{S}_R^H \mathbf{F}^H \mathbf{\Lambda}' \mathbf{F} \mathbf{S}_R}{\sigma^2} \tag{A35}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial v_x \partial v_y} \right] = \frac{1}{\sigma^2} \left\{ \begin{array}{l} \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b} + \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} \\ + \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_x} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b} \\ + \mathbf{b}^H \frac{\partial \mathbf{S}^H}{\partial v_y} (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} \end{array} \right\} \tag{A36}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial \mathbf{b}' \partial v_x} \right] = \frac{\mathbf{S}_R^H \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b} + \mathbf{S}_R^H (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_x} \mathbf{b}}{\sigma^2} \tag{A37}$$

$$-\mathbb{E} \left[\frac{\partial^2 LF_B}{\partial \mathbf{b}' \partial v_y} \right] = \frac{\mathbf{S}_R^H \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b} + \mathbf{S}_R^H (\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F}) \frac{\partial \mathbf{S}}{\partial v_y} \mathbf{b}}{\sigma^2}. \tag{A38}$$

Straightforwardly, the CRLB in case B is achieved as in (A28).

Obviously, $\mathbf{I}[\mathbf{v}]_B \succ \mathbf{I}[\mathbf{v}]_A$, where the notation \prec means that $\Delta_{AB} = \mathbf{I}[\mathbf{v}]_B - \mathbf{I}[\mathbf{v}]_A$ is a positive-definite matrix. Δ_{AB} consists of elements that satisfy $(\mathbf{F}^H \mathbf{\Lambda}' \mathbf{F})$. Straightforwardly, $\text{CRLB}_B \prec \text{CRLB}_A$. The reason for this is that any prior knowledge at the transceivers should achieve the performance gain illustrated in case B.

Another point to note is that due to the relationship between u and v in (27), both the MSE and CRLB for u are given as follows. From (27) and using $u = u_0 + \nabla u_0^T \Delta v + \mathcal{O}(\Delta v)$,

where $\nabla u_0 = \left(\begin{array}{c} j \frac{v_y}{\sqrt{1-v_y^2}} \\ \frac{-v_y}{\sqrt{1-v_y^2}} + j \frac{v_x}{(\sqrt{1-v_y^2})^3} \end{array} \right) \Big|_{v=v_0}$ and $\Delta v = \begin{pmatrix} v_x - v_{x0} \\ v_y - v_{y0} \end{pmatrix}$, ignoring higher-order infinitesimals yields

$$\text{MSE}[\Delta u] = \text{MSE}[u - u_0] \approx \nabla u_0^H \text{MSE}[\Delta v] \nabla u_0, \quad (\text{A39})$$

in which the MSE can be replaced by the CRLB for case A and case B.

References

- Li, G.; Hu, L.; Staat, P.; Elders-Boll, H.; Zenger, C.; Paar, C.; Hu, A. Reconfigurable Intelligent Surface for Physical Layer Key Generation: Constructive or Destructive? *IEEE Wirel. Commun.* **2022**, *29*, 146–153. [\[CrossRef\]](#)
- Li, G.; Sun, C.; Xu, W.; Renzo, M.D.; Hu, A. On Maximizing the Sum Secret Key Rate for Reconfigurable Intelligent Surface-Assisted Multiuser Systems. *IEEE Trans. Inf. Forensics Secur.* **2022**, *17*, 211–225. [\[CrossRef\]](#)
- Li, G.; Sun, C.; Jorswieck, E.A.; Zhang, J.; Hu, A.; Chen, Y. Sum Secret Key Rate Maximization for TDD Multi-User Massive MIMO Wireless Networks. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 968–982. [\[CrossRef\]](#)
- Danev, B.; Zanetti, D.; Capkun, S. On physical-layer identification of wireless devices. *ACM Comput. Surv. (CSUR)* **2012**, *45*, 1–29. [\[CrossRef\]](#)
- Fadul, M.K.; Reising, D.R.; Sartipi, M. Identification of OFDM-based radios under rayleigh fading using RF-DNA and deep learning. *IEEE Access* **2021**, *9*, 17100–17113. [\[CrossRef\]](#)
- Zha, X.; Chen, H.; Li, T.; Qiu, Z.; Feng, Y. Specific emitter identification based on complex Fourier neural network. *IEEE Commun. Lett.* **2021**, *26*, 592–596. [\[CrossRef\]](#)
- Tarighat, A.; Bagheri, R.; Sayed, A.H. Compensation schemes and performance analysis of IQ imbalances in OFDM receivers. *IEEE Trans. Signal Process.* **2005**, *53*, 3257–3268. [\[CrossRef\]](#)
- Tarighat, A.; Sayed, A.H. Joint compensation of transmitter and receiver impairments in OFDM systems. *IEEE Trans. Wirel. Commun.* **2007**, *6*, 240–247. [\[CrossRef\]](#)
- Kolomvakis, N.; Matthaiou, M.; Coldrey, M. IQ imbalance in multiuser systems: Channel estimation and compensation. *IEEE Trans. Commun.* **2016**, *64*, 3039–3051. [\[CrossRef\]](#)
- Chen, L.; Helmy, A.G.; Yue, G.; Li, S.; Al-Dhahir, N. Performance analysis and compensation of joint TX/RX I/Q imbalance in differential STBC-OFDM. *IEEE Trans. Veh. Technol.* **2016**, *66*, 6184–6200. [\[CrossRef\]](#)
- Chen, L.; Helmy, A.G.; Yue, G.; Li, S.; Al-Dhahir, N. Performance and compensation of I/Q imbalance in differential STBC-OFDM. In Proceedings of the 2016 IEEE Global Communications Conference (GLOBECOM), Washington, DC, USA, 4–8 December 2016; IEEE: Piscataway Township, NJ, USA, 2016; pp. 1–7.
- Zhang, X.; Li, H.; Liu, W.; Qiao, J. Iterative IQ imbalance compensation receiver for single carrier transmission. *IEEE Trans. Veh. Technol.* **2017**, *66*, 8238–8248. [\[CrossRef\]](#)
- Cheng, H.; Xia, Y.; Huang, Y.; Yang, L.; Mandic, D.P. Joint channel estimation and Tx/Rx I/Q imbalance compensation for GFDM systems. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 1304–1317. [\[CrossRef\]](#)
- Canbilen, A.E.; Ikki, S.S.; Basar, E.; Gultekin, S.S.; Develi, I. Joint impact of I/Q imbalance and imperfect CSI on SM-MIMO systems over generalized Beckmann fading channels: Optimal detection and Cramer-Rao bound. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 3034–3046. [\[CrossRef\]](#)
- Chen, Y.; You, L.; Lu, A.A.; Gao, X.; Xia, X.G. Channel estimation and robust detection for IQ imbalanced uplink massive MIMO-OFDM with adjustable phase shift pilots. *IEEE Access* **2021**, *9*, 35864–35878. [\[CrossRef\]](#)
- Chen, Y.; You, L.; Lu, A.A.; Gao, X. Widely-linear processing for the uplink of the massive MIMO with IQ imbalance: Channel estimation and data detection. *IEEE Trans. Signal Process.* **2021**, *69*, 4685–4698. [\[CrossRef\]](#)
- Cheng, X.; Luo, Z. Compensation of transmitter I/Q imbalance in millimeter-wave SC-FDE systems. *IEEE Trans. Veh. Technol.* **2016**, *66*, 4472–4476. [\[CrossRef\]](#)
- Cheng, X.; Yang, Y.; Li, S. Joint compensation of transmitter and receiver I/Q imbalances for SC-FDE systems. *IEEE Trans. Veh. Technol.* **2020**, *69*, 8483–8498. [\[CrossRef\]](#)
- Aparin, V. Analysis of CDMA signal spectral regrowth and waveform quality. *IEEE Trans. Microw. Theory Tech.* **2001**, *49*, 2306–2314. [\[CrossRef\]](#)

20. Gard, K.G.; Larson, L.E.; Steer, M.B. The impact of RF front-end characteristics on the spectral regrowth of communications signals. *IEEE Trans. Microw. Theory Tech.* **2005**, *53*, 2179–2186. [[CrossRef](#)]
21. Liu, M.W.; Doherty, J.F. Nonlinearity estimation for specific emitter identification in multipath channels. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 1076–1085.
22. Pan, Y.; Yang, S.; Peng, H.; Li, T.; Wang, W. Specific emitter identification based on deep residual networks. *IEEE Access* **2019**, *7*, 54425–54434. [[CrossRef](#)]
23. Liu, W.; Li, X.; Huang, Z.; Wang, X. Transmitter fingerprinting for VLC systems via deep feature separation network. *IEEE Photonics J.* **2021**, *13*, 1–7. [[CrossRef](#)]
24. Li, J.; Matthaiou, M.; Svensson, T. I/Q imbalance in two-way AF relaying. *IEEE Trans. Commun.* **2014**, *62*, 2271–2285. [[CrossRef](#)]
25. Rappaport, T.S. *Wireless Communications: Principles and Practice, 2nd ed.*; Pearson Education: Upper Saddle River, NJ, USA, 2010.
26. Lecomte, C. Exact statistics of systems with uncertainties: An analytical theory of rank-one stochastic dynamic systems. *J. Sound Vib.* **2013**, *332*, 2750–2776. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.