



# Article Digital Self-Interference Canceler with Joint Channel Estimator for Simultaneous Transmit and Receive System

Shiyu Song<sup>1</sup>, Yanqun Tang<sup>1,\*</sup>, Xianjie Lu<sup>1</sup>, Yu Zhou<sup>1</sup>, Xizhang Wei<sup>1</sup>, Zhengpeng Wang<sup>1</sup> and Songhu Ge<sup>2</sup>

- <sup>1</sup> School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China; songshy7@mail2.sysu.edu.cn (S.S.); luxj25@mail2.sysu.edu.cn (X.L.); zhouy633@mail2.sysu.edu.cn (Y.Z.); weixzh7@mail.sysu.edu.cn (X.W.); wangzhp26@mail2.sysu.edu.cn (Z.W.)
- <sup>2</sup> National Key Laboratory of Electromagnetic Energy, Naval University of Engineering, Wuhan 430033, China; gesonghu@126.com
- \* Correspondence: tangyq8@mail.sysu.edu.cn

Abstract: Simultaneous transmit and receive wireless communications have been highlighted for their potential to double the spectral efficiency. However, it is necessary to mitigate self-interference (SI). Considering both the SI channel and remote transmission (RT) channel need to be estimated before equalizing the received signal, we propose two adaptive algorithms for linear and nonlinear self-interference cancellation (SIC), based on a multi-layered joint channel estimator structure. The proposed algorithms estimate the RT channel while performing SIC, and the multi-layered structure ensures improved performance across various interference-to-signal ratios. The M-estimate function enhances the robustness of the algorithm, allowing it to converge even when affected by impulsive noise. For nonlinear SIC, this paper introduces an adaptive algorithm based on generalized Hammerstein polynomial basis functions. The simulation results indicate that this approach achieves a better convergence speed and normalized mean squared difference compared to existing SIC methods, leading to a lower system bit error rate.

**Keywords:** simultaneous transmit and receive; self-interference cancellation; digital cancellation; adaptive filtering; joint channel estimator

# 1. Introduction

Full duplex (FD) relay technology enables a communication system to simultaneously transmit and receive signals on the same frequency through a relay node [1]. Inspired by this technology, simultaneous transmit and receive (STAR) technology, also referred to as in-band full-duplex (IBFD), has emerged as a novel innovation within the field. It aims to allow communication users to directly transmit and receive signals at the same carrier frequency, facilitating an even more efficient use of spectral resources. Moreover, STAR technology can be synergistically combined with other advanced technologies that enhance spectral utilization, such as reconfigurable intelligent surfaces [2], millimeter-wave technology [3], and multiple access techniques [4]. As a result, STAR technology has arisen as a promising approach to increase data throughput and solve the problem of scarce spectrum resources, which can be applied in the field of wireless communications, radar, satellites, and unmanned aerial vehicles [5,6].

However, the self-interference (SI) signal, which is generated from the near end, can be 100–120 dB stronger than the weak received signal, thus making the receiver inoperable [7]. Therefore, self-interference cancellation (SIC) technology is key to the implementation of STAR communications. At present, the SI signal can be eliminated in several stages and various domains, as shown in Figure 1. Typically, SIC is achieved through the cascaded antenna domain [8], analog domain [9], and digital domain [10–15]. The antenna domain SIC is passive in nature, and mainly uses antenna isolation or related beamforming algorithms to prevent RF reception path blockage. In the analog domain, directly coupled



**Citation:** Song, S.; Tang, Y.; Lu, X.; Zhou, Y.; Wei, X.; Wang, Z.; Ge, S. Digital Self-Interference Canceler with Joint Channel Estimator for Simultaneous Transmit and Receive System. *Sensors* **2024**, *24*, 2449. https://doi.org/10.3390/s24082449

Academic Editor: Stefano Savazzi

Received: 17 March 2024 Revised: 6 April 2024 Accepted: 8 April 2024 Published: 11 April 2024



**Copyright:** © 2024 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/). interference suppression and digital-assisted analog interference-suppression methods are commonly used at the receiver low-noise amplifier (LNA) input before the analog-to-digital conversion. The last stage is digital SIC, which aims to eliminate the remaining SI from the received signal by reconstructing the SI signal in the receiving link and reducing its power below the noise floor. These steps have been applied in some prototypes and have demonstrated the potential of STAR communications [7,16]. The remainder of this paper focuses on digital SIC algorithm design and both the linear and nonlinear SIC algorithms are evaluated through extensive simulations, demonstrating their effectiveness in improving the performance of STAR communication systems.



Figure 1. A block diagram illustrating the STAR system and various forms of SIC.

Digital SIC has attracted significant attention in recent years to address the SI problem in STAR communication systems. Nonetheless, the practical application of digital SIC still faces many challenges. First of all, the observed SI signal exhibits significant frequency and time selectivity due to reflections within the device and its surrounding environment, which demonstrate temporal variations. This can be regarded as a classic system identification problem, which is typically solved through adaptive finite impulse response (FIR) filters. Secondly, in practical communication environments, there are often additional sources of noise, such as electromagnetic interference, lightning noise, radar signals, and other human-made disturbances. These types of noise typically exhibit strong pulse-like characteristics within very short time durations, which can lead to performance degradation of various filtering algorithms based on the assumption of Gaussian noise. Furthermore, the significantly higher power of the SI signal compared to the desired signal at the receiver implies that even minor distortions can significantly degrade the signal of interest. These hardware impairments, such as phase noise, in-phase/quadrature (I/Q) imbalance, and power amplifier (PA) and baseband nonlinearity, among others, with PA nonlinearity are especially harmful to the digital canceler, making digital SIC insufficient.

Digital cancellation algorithms accounting for impulsive noise have been proposed in earlier research [10,17]. The work in [11,12] focused on linear SIC by employing the least mean square (LMS) adaptive filter. In addition, different linear SIC methods were compared in [13]. To compensate for nonlinear PA distortion, the authors in [18] proposed novel predistorters and their parameter extraction algorithms. The proposed self-adaptive nonlinear digital cancelers in [14,15] both utilize a novel orthogonalization procedure for nonlinear basis functions, together with low-cost LMS-based parameter learning. Both PA nonlinearity and I/Q imbalances were considered in [19,20], and more impairments, including local oscillator leakage and baseband nonlinearity, were analyzed and measured on the Universal Software Radio Peripheral in [7]. Some studies have also proposed various frequency domain SIC methods [21,22]. In [21], a frequency domain SI canceler based on the parallel Hammerstein (PH) model for an IBFD orthogonal frequency division multiplexing System was proposed, while Ref. [22] utilized a successive cancellation cascaded structure as a replacement for the adaptive filter, effectively reducing the computational load of SIC processing and simplifying the procedure. Most published works are based on the scenario where only the own transmitter (TX) signal is known, without considering the remote transmission (RT) signal. When the RT sends a signal for channel estimation, it can be assumed to be known, and additional prior information will further enhance the

performance of SIC [23]. In this paper, a robust multi-layered M-estimate total least mean square (m-MTLS) canceler is proposed to enhance the performance in scenarios where both the data matrix and the observation signal are also affected by impulsive noise. The method exhibits significant robustness against impulse noise, managing to maintain a low convergence error even in the presence of impulse noise. When both the PA nonlinearities of the local transmitter and the training signal from RT are taken into account, we introduce a novel algorithm for nonlinear digital SIC based on a set of generalized Hammerstein polynomials (HPs) [15], by estimating the SI channel and the RT channel at the same time. This method of joint estimation can more accurately estimate the remote channel while reducing the residual SI signal, compared with existing SIC methods. The simulation results show that this method exhibits a faster convergence rate and demonstrably enhances system performance as the signal-to-noise ratio (SNR) increases.

The rest of this paper is organized as follows. In Section 2, the basic STAR system and linear SI signal modeling are first described. Then, considering the nonlinearity introduced by the PA, the SI model is extended to a nonlinear model, followed by the introduction of a joint estimator that can simultaneously estimate both SI and RT channels. Section 3 proposes two new algorithms for linear and nonlinear SIC based on adaptive FIR, respectively. The linear SIC also enhances the robustness of the algorithm under impulsive noise interference. The simulation results are reported in Section 4. Finally, Section 5 concludes the paper.

In the following, continuous and discrete time signals are expressed in italic lowercase, while vectors are denoted by bold lowercase symbols. The acronyms used throughout the paper are also summarized in Table 1.

Acronym	Definition		
STAR	simultaneous transmit and receive		
IBFD	in-band full-duplex		
SI	self-interference		
SIC	self-interference cancellation		
RF	Radio Frequency		
LNA	low-noise amplifier		
PA	power amplifier		
LMS	least mean square		
PH	parallel Hammerstein		
TX	transmitter		
RT	remote transmitter		
HP	Hammerstein polynomials		
SNR	signal-to-noise ratio		
ISR	interference-to-signal ratio		
m-MTLS	multi-layered M-estimate total least squares		
NMSD	normalized mean squared difference		
MSE	mean squared error		
BER	bit error rate		
DFE	decision feedback equalizer		

Table 1. Table of acronyms.

### 2. System Model

The structure of the analyzed STAR system is presented in Figure 2, with signals propagating at the different stages. We denote the baseband transmission at time slot n as x[n], which is perfectly known. In this analysis, the RF components are assumed to be ideal and the nonlinear model will be discussed later. The received signal, after being processed

by the LNA and subjected to SIC in the analog domain, is converted into a baseband signal by the ADC at time slot *n* and can be written as

$$y[n] = r[n] + f(x[n], x[n-1], \dots x[n-L+1]) + v[n]$$
  
= r[n] + s[n] + v[n], (1)

where r[n] and v[n] denote the desired signal from the remote source and noise at time instant n, respectively. The received signal is assumed to be generated by a function  $f(\cdot)$  with L previous transmitted signals. Furthermore, the SI signal from the local TX, s[n], can be represented by a linear model as

$$s[n] = f(x[n], x[n-1], \dots, x[n-L+1]) = \boldsymbol{w}^{T}[n]x_{L}[n],$$
(2)

where  $w \in \mathbb{R}^{N \times 1}$  denotes the SI channel and  $x_L[n] = [x[n], x[n-1], \dots, x[n+L-1]]$ .



Figure 2. The signals propagating at the different stages of the analyzed STAR system.

When the nonlinear hardware impairments are taken into account, one particularly important imperfection is PA nonlinearity; thus, we assume that the transmitted signal is only distorted by high-order harmonics of the PA, excluding the effects of I/Q imbalance or phase noise. The PAs exhibit nonlinearity when power-efficient operation is sought, and it is commonly modeled with the well-known PH model [14,19,22,24]. With an input vector  $x_L[n]$ , the output of a *P*-th order model can be expressed as

$$x_{\rm PA}(t) = \sum_{\substack{p=1\\p \text{ odd}}}^{\frac{p+1}{2}} h_{p,\rm PA}(t) * \phi_p(x(t)),$$
(3)

where *P* is the highest nonlinearity order of the model,  $h_{p,PA}(t)$  represents the response, and  $\phi_p(x(t)) = |x(t)|^{2p}x(t)$  is the *p*th-order basis function. The symbol \* represents the convolution operation. We assume that the PA memory length is  $L_1$  and the length of the SI channel is  $L_2$ . Then, the observed SI signal passes through the SI channel and can be rewritten as

$$s[n] = \sum_{\substack{p=1\\p \text{ odd}}}^{\frac{p+1}{2}} \sum_{l=0}^{L-1} w_p[l] \phi_p(x[n-l]),$$
(4)

where  $w_p$  is the response weight of the entire channel, with a length of  $L = L_1 + L_2 - 1$ .

Building on the analysis presented earlier, for the purpose of mitigating residual SI within the digital domain, it is imperative to acquire the optimal estimate of the SI channel response, denoted as  $\hat{w}[n]$ , corresponding to w[n]. This involves subtracting

the reconstructed SI signal from y[n]. Followed by the SIC process, the remainder of the signal is presumed to embody the RT signal. Subsequent stages may include RT channel estimation and equalization, facilitating the detection of data symbols embedded within the RT signal.

From the SI model mentioned in (2) and (4), it is evident that the performance of digital SIC is fundamentally contingent upon the accuracy of the channel response coefficient estimation. Moreover, the power of the residual SI signal remains high after RF cancellation. when the interference-to-signal ratio (ISR) is high, even minimal residuals can severely compromise the separation of the RT signal. Additionally, at higher SNRs, noise ceases to be the predominant factor impacting the estimation of the SI signal. Under such circumstances, the primary source of error in SIC emanates from the RT signal itself. Consequently, in scenarios where the RT signal is known, it becomes essential to employ a joint channel estimator to enhance the precision of both SI and RT signal separations, thereby mitigating the aforementioned challenges. We assume that the original RT signal, i[n], is known; then, (1) and (2) can be rewritten in [23] as

$$y[n] = \boldsymbol{w}^{T}[n]\boldsymbol{x}[n] + \boldsymbol{h}^{T}[n]\boldsymbol{i}[n] + \boldsymbol{v}[n],$$
(5)

where h[n] denotes the RT channel coefficient, and the interest signal, r[n] is represented as  $h^{T}[n]i[n]$ .

By merging two channels into a joint channel, according to (5), we can obtain the following expression:

$$y[n] = \boldsymbol{c}^{T}[n]\boldsymbol{u}[n] + v[n], \tag{6}$$

where  $c^{T}[n] = [w^{T}[n], h^{T}[n]]$  combines the coefficients of both the SI and RT channels, and  $u[n] = [x^{T}[n], i^{T}[n]]^{T}$  is the new input signal.

In (5), the estimation of SI is influenced by both  $r[n] = h^T[n]i[n]$  and v[n]. However, by combining the SI and RT channel as a new estimation parameter, only the noise v[n] impacts this process. The benefits of this joint channel estimation approach will be detailed in the simulations presented in Section 4.

#### 3. Proposed Digital Canceler

When considering the SI and RT channels, it is important to acknowledge that they will vary over time, as the environmental reflection paths are also time-variant. Therefore, we opted for adaptive filtering algorithms and incorporated a multi-layered structure, as shown in Figure 3, to enhance the RT estimation performance of the joint estimator.



Figure 3. The multi-layered structure of the joint channel estimator.

As depicted in Figure 3, the estimation of joint channel coefficients  $\hat{c}[n]$  in the first layer is given by  $\hat{c}_{(1)}^{T}[n] = [\hat{w}_{(1)}^{T}[n], \hat{h}_{(1)}^{T}[n]]$ ; then, the residual signal  $y_{(2)}[n]$ , obtained by subtracting the reconstructed SI signal from y[n], serves as the input for the second layer. Furthermore, we can obtain

$$y_{(2)}[n] = c_{(1)}^{T}[n]u[n] - \hat{w}_{(1)}^{T}[n]x[n] + v[n]$$
  
=  $c_{(2)}^{T}[n]u[n] + v[n].$  (7)

This procedure is consistently implemented across successive layers of the filter. In the *m*-th layer, this process is articulated as

$$y_{(m+1)}[n] = y_{(m)}[n] - \hat{w}_{(m)}^{T}[n]\mathbf{x}[n] + v[n] = \mathbf{c}_{(m+1)}^{T}[n]\mathbf{u}[n] + v[n]$$
(8)

where  $y_{(m)}[n]$  is the input of the *m*-th layer and  $\hat{w}_{(m)}[n]$  is the estimation of the SI channel, and we have  $\hat{c}_{(m)}^{T}[n] = [\hat{w}_{(m)}^{T}[n], \hat{h}_{(m)}^{T}[n]]$  as the estimation of the joint channel.

For each layer, while subtracting the residual from the received signal, an estimation of the RT channel  $\hat{h}_{(m)}[n]$  is performed and the output of the last layer  $y_{(M+1)}[n]$  is regarded as the interest signal  $\hat{r}[n]$ .

#### 3.1. Robust m-MTLS Adaptive Algorithm

We consider a linear model where both the independent and dependent variables are subject to measurement errors, which can be illustrated by

$$(y_n + v_n) = (x_n + u_n)^T \boldsymbol{w}, \tag{9}$$

where  $w \in \mathbb{R}^{L \times 1}$  denotes the system vector to be estimated, with  $x_n \in \mathbb{R}^{L \times 1}$  as the input vector and  $y_n \in \mathbb{R}$  as the output signal at time *n*. The noise vectors  $u_n$ , distributed as  $N(\mathbf{0}, \sigma_i^2 \mathbf{I}) \in \mathbb{R}^{L \times 1}$ , represent the input noise, and  $v_n$ , distributed as  $N(\mathbf{0}, \sigma_o^2) \in \mathbb{R}$ , represent the output noise, where  $\sigma_i^2$  and  $\sigma_o^2$  are the variances in the input and output noise, respectively. Here, **0** indicates the zero vector and **I** the identity matrix.

Total least squares (TLS) considers errors in all variables, making it particularly useful for more accurate modeling and estimation when errors are present in both predictors and outcomes [25]. Its cost function can be formulated as

$$\min_{\boldsymbol{w}} J(\boldsymbol{w}) = \frac{1}{N} \sum_{n=1}^{N} \frac{\left(\tilde{y}_n - \boldsymbol{w}^T \tilde{\boldsymbol{x}}_n\right)^2}{\|\boldsymbol{w}\|^2 + \gamma},$$
(10)

where  $\gamma \stackrel{\text{def}}{=} \sigma_o^2 / \sigma_i^2$  is a parameter to normalize the noise variances,  $\tilde{y}_n = y_n + v_n$  and  $\tilde{x}_n = x_n + u_n$ .

Similar to other adaptive algorithms, the iterative formula for TLS based on stochastic gradient descent can be expressed as

$$\hat{\boldsymbol{w}}_{n+1} = \hat{\boldsymbol{w}}_n - \mu \hat{\boldsymbol{g}}(\hat{\boldsymbol{w}}_n) = \hat{\boldsymbol{w}}_n + \mu \alpha_n (\tilde{\boldsymbol{x}}_n + \alpha_n \hat{\boldsymbol{w}}_n),$$
(11)

where  $\mu$  is the step-size parameter, and  $\hat{g}$  is the instantaneous gradient of J(w), which can be represented as

$$\hat{g}(\boldsymbol{w}) = \alpha_n (\tilde{\boldsymbol{x}}_n + \alpha_n \boldsymbol{w}_n),$$

$$\alpha_n \stackrel{\text{def}}{=} \frac{\tilde{y}_n - \tilde{\boldsymbol{x}}_n^T \boldsymbol{w}_n}{\|\boldsymbol{w}_n\|^2 + \gamma}.$$
(12)

To enhance the robustness of the algorithm when the received signal is subjected to impulsive noise, an M-estimate function can be utilized to improve TLS. This improved algorithm, referred to as the M-estimate total least mean square (MTLS) adaptive algorithm, has its cost function defined as

$$\min_{\boldsymbol{w}} J(\boldsymbol{w}) = E\left[\frac{\rho(e_n^2)}{\|\boldsymbol{w}\|^2 + \gamma}\right],\tag{13}$$

where  $\rho(\cdot)$  is the M-estimate function, which is a real-valued even function, given by

$$\rho(e_n) = \begin{cases} e_n^2/2, & |e_n| < \xi \\ \xi^2/2, & |e_n| \ge \xi, \end{cases}$$
(14)

where  $\xi = c_1 \hat{\sigma}_e$  represents the threshold parameter, with  $c_1$  set to 2.576, and  $\hat{\sigma}_e$  can be calculated by

$$\hat{\sigma}_{e_n}^2 = \lambda_\sigma \hat{\sigma}_{e_{n-1}}^2 + c_2(1 - \lambda_\sigma) \operatorname{med}(A_{e_n}), \tag{15}$$

where  $A_e[n] = [e^2[n], e^2[n-1], \dots, e^2[n-N_w+1]]$ , med(·) denotes the median operation, and the parameter  $\lambda_\sigma$ , indicative of the weighting factor, is typically selected within the range of 0.98 to 0.99. The constant  $c_2 = 1.483(1+5/(N_w-1))$ , and  $N_w$  represents the length of the window over which the estimation is performed.

The iterative formula (11) can be rewritten as

$$\hat{g}(\boldsymbol{w}) = \begin{cases} -\frac{\left(\|\boldsymbol{w}\|^2 + \gamma\right)e_n\tilde{x}_n + e_n^2\boldsymbol{w}}{\left(\|\boldsymbol{w}\|^2 + \gamma\right)^2}, & |\boldsymbol{e}_n| < \xi\\ 0, & |\boldsymbol{e}_n| \ge \xi, \end{cases}$$
(16)

and

$$\hat{\boldsymbol{w}}_{n+1} = \hat{\boldsymbol{w}}_n - \mu \hat{\boldsymbol{g}}(\hat{\boldsymbol{w}}_n). \tag{17}$$

By integrating the MTLS adaptive algorithm with the multi-layered joint channel estimator discussed in Section 2, a robust linear SI canceler can be achieved [10].

#### 3.2. Multi-Layered Generalized HP-Based Adaptive Algorithm

Nonlinear SIC can be regarded as a problem of nonlinear system identification. Before engaging in adaptive filtering, it is necessary to linearize the SI signal using the basis functions described in (4).

The nonlinear PA model first proposed in [15] employs a set of generalized HPs for its representation. The p-th order nonlinear basis function in Equation (3) is defined as

$$\phi_p(x(t); c_p) = \sum_{k=0}^{p-1} c_{p,k} |x(t)|^{2k} x(t),$$
(18)

which can be rewritten in the digital domain as

$$\phi_p(x[n]; c_p) = \sum_{k=0}^{p-1} c_{p,k} |x[n]|^{2k} x[n], \qquad (19)$$

where  $c_p = [c_{p,0}, \dots c_{p,p-1}]^T$  represents the polynomial coefficients of the *p*-th basis function. When the polynomial coefficients satisfy  $c_{p,p-1} = 1$  and  $c_{p,k} = 0, \forall k \neq p-1$ , the basis

function simplifies to the standard HP basis function, i.e.,  $\phi_p(x[n]) = |x[n]|^{p-1}x[n]$ . Then, the SI signal s[n] can be represented by the aforementioned generalized HP as

$$s[n] = \sum_{p=1}^{\frac{p+1}{2}} \sum_{l=0}^{L-1} w_p[l] \phi_p(x[n-l]; c_p)$$

$$= \sum_{p=1}^{\frac{p+1}{2}} \sum_{l=0}^{L-1} w_p[l] \left( \sum_{k=0}^{p-1} c_{p,k} |x[n-l]|^{2k} x[n-l] \right).$$
(20)

To enhance the convergence speed of the adaptive filtering algorithm, it is also necessary to determine the coefficients  $c_p = [c_{p,0}, \dots c_{p,p-1}]^T$  of the generalized HP basis

$$\begin{cases} E[\phi_p(x[n]; c_p)\phi_j(x[n]; c_j)] = 1, p = j \\ E[\phi_p(x[n]; c_p)\phi_j(x[n]; c_j)] = 0, p \neq j \end{cases}$$
(21)

For polynomial basis functions with the highest order term of 2p - 1 (p > 1), assuming orthogonality with any basis function of order less than 2p - 1, it is possible to derive p - 1 orthogonal equations as

$$E[\phi_p(x[n]; c_p)\phi_k(x[n]; c_k)] = 0, k = 1, 2, \dots, p-1.$$
(22)

Furthermore, we can obtain

$$\begin{cases} E[\phi_p(x[n]; c_p) c_{1,0} x[n]] = 0, (k = 1) \\ E[\phi_p(x[n]; c_p) (c_{2,0} x[n] + c_{2,1} |x[n]|^2 x[n]) ] = 0, (k = 2) \\ \vdots \\ E[\phi_p(x[n]; c_p) (c_{p-1,0} x[n] + \dots + c_{p-1,p-2} |x[n]|^{2p-2} x[n]) ] = 0, (k = p-1). \end{cases}$$
(23)

Assume  $c_{p,p-1} = 1, \forall p$ . When k = 1, we define  $\mu_p = E[(x[n])^p]$ ; then, (23) can be rewritten as

$$E\left[\left(c_{p,0}x[n] + c_{p,1}|x[n]|^{2}x[n] + \dots + |x[n]|^{2p-2}x[n]\right)x[n]\right]$$
  
=  $E\left[\left(c_{p,0}|x[n]|^{2} + c_{p,1}|x[n]|^{4} + \dots + |x[n]|^{2p}\right)\right]$ , (k = 1). (24)  
=  $c_{p,0}\mu_{2} + c_{p,1}\mu_{4} + \dots + \mu_{2p}$   
= 0

Let  $\mu_{2a}^{2b} = {\mu_{2a}, \mu_{2a+2}, \dots, \mu_{2b}}, 1 \le a \le b$ . Equation (23) can be written in matrix form as

$$\mathbf{C}_{p}\mathbf{M}_{p}\bar{\mathbf{c}}_{p} + \mathbf{C}_{p}\boldsymbol{\mu}_{2p}^{4p-4} = 0, \tag{25}$$

where  $\bar{c}_p$  is a subvector of the  $c_p = \left[\bar{c}_p^T \ 1\right]^T$  and  $\mathbf{C}_p$  is an invertible lower triangular matrix. The matrix  $\mathbf{M}_p$  is defined as

$$\mathbf{M}_{p} = \begin{bmatrix} \mu_{2} & \mu_{4} & \cdots & \mu_{2p-2} \\ \mu_{4} & \mu_{6} & \cdots & \mu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{2p-2} & \mu_{2p} & \cdots & \mu_{4p-6} \end{bmatrix} \text{ and } \det(\mathbf{M}_{p}) \neq 0$$
(26)

Thus, we have

$$\bar{c}_p = -\mathbf{M}_p^{-1} \mu_{2p}^{4p-4}.$$
(27)

The instantaneous basis function vector is defined as

$$\boldsymbol{\Phi}[n] = \left[\boldsymbol{\phi}_1(\boldsymbol{x}_L[n]; \boldsymbol{c}_1), \dots, \boldsymbol{\phi}_{\frac{P+1}{2}}\left(\boldsymbol{x}_L[n]; \boldsymbol{c}_{\frac{P+1}{2}}\right)\right]^T.$$
(28)

Integrating the nonlinear PA model with the multi-layered joint channel estimator employing the LMS algorithm results in

$$y[n] = \boldsymbol{w}^{T}[n]\boldsymbol{\Phi}[n] + \boldsymbol{h}^{T}[n]\boldsymbol{i}[n] + \boldsymbol{v}[n]$$
  
=  $\begin{bmatrix} \boldsymbol{w}^{T}[n] \ \boldsymbol{h}^{T}[n] \end{bmatrix} \begin{bmatrix} \boldsymbol{\Phi}^{T}[n] \ \boldsymbol{i}^{T}[n] \end{bmatrix}^{T} + \boldsymbol{v}[n]$   
=  $\boldsymbol{h}^{T}_{joint}[n]\boldsymbol{u}[n] + \boldsymbol{v}[n],$  (29)

where  $\boldsymbol{h}_{joint}^{T}[n] = \begin{bmatrix} \boldsymbol{w}^{T}[n] & \boldsymbol{h}^{T}[n] \end{bmatrix}$  and  $\boldsymbol{u}[n] = \begin{bmatrix} \Phi^{T}[n] & \boldsymbol{i}^{T}[n] \end{bmatrix}^{T}$ . Furthermore, the iterative formula can be expressed as

$$\hat{\boldsymbol{h}}_{joint}[n+1] = \hat{\boldsymbol{h}}_{joint}[n] + \mu \boldsymbol{u}[n]\boldsymbol{e}[n], \tag{30}$$

where  $\mu$  is the step size.

At the *m*-th layer, the received signal after digital SIC is

$$y_{(m+1)}[n] = y_{(m)}[n] - \hat{\boldsymbol{h}}_{joint(m)}^{T} [n] \Phi[n] + v[n].$$
(31)

The *M*-layered adaptive nonlinear SIC algorithm is summarized in Algorithm 1 and Table 2.

#### Algorithm 1 Adaptive algorithm for nonlinear SIC

**Input:** received signal y[n], known signals x[n] and i[n], basis order *P*, moment collect length  $N_m$  and basis function generation interval  $N_{int}$ 

**Output:** signal after SIC  $\hat{r}[n]$  and RT channel estimation  $\hat{h}_{(m)}^{1}[n]$ 

**Initialization:**  $n_s = 1$  and  $\boldsymbol{h}_{joint(m)}^T[n] = \boldsymbol{0}$ 1: for n = 1, 2, ..., N do if  $n = n_s + N_m - 1$  then  $\mu_p = \frac{1}{N} \sum_{n_s}^{n_s + N_m - 1} (x[n])^p$ for  $p = 2, 3, \dots, \frac{p+1}{2}$  do  $\bar{c}_p = -\mathbf{M}_p^{-1} \mu_{2p}^{4p-4}$ , basis function  $\phi_p(x[n]; c_p)$  generation 2: 3: 4: 5: end for 6: end if 7: if  $n = n_s + N_{int} - 1$  then 8:  $n_s = n$ 9: 10: end if end if  $y_{(1)}[n] = y[n] = h_{joint}^{T}[n]u[n] + v[n]$ for m = 1, 2, ..., M do  $e_{(m)}[n] = y_{(m)}[n] - \hat{h}_{joint(m)}^{T}[n]u[n]$   $\hat{h}_{joint(m)}^{T}[n+1] = \hat{h}_{joint(m)}^{T}[n] + \mu u[n]e_{(m)}[n]$   $\hat{w}_{(m)}^{T}[n+1] = (\hat{h}_{joint(m)}^{T}[n+1])_{1:N_{1}}$ 11: 12: 13: 14: 15:  $\hat{\boldsymbol{h}}_{(m)}^{T}[n+1] = \left(\hat{\boldsymbol{h}}_{joint(m)}^{T}[n]\right)_{N_{1}+1:N_{1}+N_{2}}$ 16:  $y_{(m+1)}[n] = y_{(m)}[n] - \hat{w}_{(m)}^T[n]x[n]$ 17: end for 18: 19:  $\hat{r}[n] = y_{(m+1)}[n]$ 20: end for

Parameter	Symbol	
<i>p</i> -th order basis functions	$\phi_p(x[n]; \boldsymbol{c}_p)$	
Coefficients of the <i>p</i> -th basis function	$c_p$	
Nonlinear order of PA	P	
<i>p</i> -th SI channel response	$w_p$	
<i>p</i> -th RT channel response	$h_{v}$	
Joint channel response	$h_{ioint}$	
<i>p</i> -th even-order partial moments	$\mathbf{M}_{p}$	
Number of layers	Ń	
Input of <i>m</i> -th layers	$y_{(m)}$	

#### Table 2. Table of notations.

# 4. Simulation Results

4.1. Linear SI Canceler

In the scenario described by Equation (9), the unknown vector w, with a dimensionality of L = 14, conforms to the condition  $||w||^2 = 1$ . Additionally, the elements of ware distributed according to a Gaussian distribution with a mean of zero. The input signal is independently generated from zero-mean Gaussian with unit variance. The input and output noises are denoted as  $v_{in} = v_b$  and  $v_{out} = v_a + v_i$ , respectively, where  $v_a$ and  $v_b$  are Gaussian-distributed with equal variances of  $\sigma_a^2 = \sigma_b^2 = 0.1$ . The Bernoulli– Gaussian (BG) process is utilized as an impulsive noise model. The impulsive noise  $v_i[n] = b[n]p[n]$ , where b[n] is a Bernoulli process with the probability density function  $P(b[n] = 1) = P_i, P(b[n] = 0) = 1 - P_i$  and  $p[n] \sim N(0, 10)$  [26]. Their performance is displayed in Figure 4, quantitatively evaluated via the normalized mean squared difference (NMSD), defined as  $= 10 \log(||\hat{w}_n - w||^2 / ||w||^2)$ . It can be observed that the LMS and TLS algorithms decline due to noise interference, whereas the MTLS algorithm remains largely unaffected by such disturbances.



**Figure 4.** Average NMSD learning curve with different impulsive noise probabilities. (a)  $P_i = 0$ ; (b)  $P_i = 0.01$ .

We consider a baseband STAR system characterized by a 5 MHz signal bandwidth. The system operates at a sampling rate of 10 MHz, with both the local and remote transmitters employing binary phase shift keying modulation for signal transmission. The SI and RT channels have lengths specified as  $N_1 = 4$  and  $N_2 = 10$ , respectively. Furthermore, their average energy follows the conditions ISR =  $E(||w||^2/||h||^2)$  and SNR =  $E(||s(n)||^2/\sigma_a^2)$ . The received signal y(n) is subject to corruption from two sources: Gaussian white noise

 $v_a$  and random impulsive noise  $v_i$ . The occurrence probability of the impulsive noise is denoted as  $P_i$ . Additionally, the local reference signal i(n) is influenced by Gaussian noise  $v_b$ , where the condition  $\sigma_a^2 = \sigma_b^2$  holds true. In the context of this simulation, we employ the m-MTLS algorithm with a total of number of M = 2 layers, which is mainly a compromise considering the additional computational cost brought by multi-layer filters. As M increases, the effectiveness of SIC initially improves and then tends to stabilize or even decline. This is due to the propagation of estimation errors across different layers. To assess the efficacy of the proposed multi-layered joint estimator, a comparative analysis is conducted against the minimum mean square error (MMSE) estimator and the single-layer MTLS estimator, focusing on their performance in estimating the RT channel. The physical meaning of the parameters involved in this simulation is summarized in Table 3.

Table 3.	Table	of	parameters.
----------	-------	----	-------------

Parameter	Symbol	Value
Length of SI channel	$N_1$	4
Length of RT channel	$N_2$	10
Layers	M	2
Input noise	$v_{in}$	$v_b$
Output noise	v <sub>out</sub>	$v_a + v_i$
Gaussian noise	$v_a, v_b$	$\sigma_a^2 = \sigma_b^2$
Impulsive noise	$v_i$	$\sigma_{i}^{2} = 10$
ISR	-	20 dB

Figure 5 illustrates the performance curves of several algorithms in terms of NMSD as the SNR varies across different noise environments, with ISR = 20 dB. Compared to the MMSE estimator, the proposed method exhibits higher robustness. The MMSE estimator shows a noticeable decline in performance at certain sampling points when affected by impulsive noise pollution. However, thanks to the M-estimate function, the m-MTLS estimator remains virtually unaffected. Furthermore, as the SNR increases, the NMSD of the m-MTLS estimator gradually decreases, whereas the MMSE estimator experiences limited gain due to convergence being disrupted by interference from the interest signal.



**Figure 5.** The performance of different estimators under the impulse noise probability. (a)  $P_i = 0.01$ ; (b)  $P_i = 0.05$ .

In this section, We use the Saleh's PA model [27] to simulate the nonlinear distortion, and the SI signal received after ADC can be described as follows:

$$f(x[n]) = \sum_{l=0}^{L-1} h[l] \frac{\gamma x[n-l]}{1+\beta |x[n-l]|^2},$$
(32)

where  $\gamma = 3$  and  $\beta = 0.03$  in this simulation, and h[n] is the overall memory length, set to L = 5.

We assess the performance in terms of the mean squared error (MSE), defined as follows:

$$MSE = \mathbb{E}\Big[ (f(\boldsymbol{x}[n]) - \hat{\boldsymbol{y}}[n])^2 \Big].$$
(33)

The performance of the proposed algorithm and the commonly used LMS algorithms based on HPs are evaluated and the results are shown in Figure 6. In this simulation, we assume the transmitted data follow a uniform distribution in the range [-2.5, 2.5] and v[n] is zero-mean Gaussian noise with variance 1.

As shown in Figure 6, compared to existing algorithms based on standard HPs, the proposed algorithm achieves faster convergence while maintaining equivalent MSE performance.



**Figure 6.** MSE performance of different SIC algorithms with the same step size, when P = 5.

In the nonlinear SIC scenario, we simulate data transmission using a uniform distribution within the range of [-2.5, 2.5], and the remote signal  $i[n] \in \{-1, 1\}$ . The nonlinear SI canceler is demonstrated in Section 3.2. Considering both computational complexity and algorithm performance, the joint SIC algorithm has a two-layer filtering structure with M = 2. The basis functions for all algorithms are of order up to five.

Figures 7 and 8 compare the power spectrum of different signals at the receiver for ISR = 20 dB and ISR = 40 dB, respectively. As seen, in both cases, the signal after SIC more closely approximates the remote signal when using the proposed SIC algorithm compared to the method based on HP-LMS.

Figure 9 shows the NMSD and bit error rate (BER) performance of the proposed algorithm and other SIC algorithms. For symbol detection at the receiver, a decision feedback equalizer (DFE) is employed, with both the feedforward and feedback filters having lengths of 15. The DFE employs a truncated version of the transmitted signal as the training signal, with tap weights being trained exclusively during the first block. In this result, as expected, the joint SIC algorithm with a generalized HP basis function with orthogonality exhibits superior performance.



**Figure 7.** Power spectrum of signals at the receiver, with ISR = 20 dB. (a) Proposed algorithm; (b) HP-LMS algorithm.



**Figure 8.** Power spectrum of signals at the receiver, with ISR = 40 dB. (a) Proposed algorithm; (b) HP-LMS algorithm.



**Figure 9.** Performance of different algorithms, with ISR = 40 dB. (a) NMSD performance; (b) BER performance.

# 5. Conclusions

In this paper, we proposed two adaptive algorithms for linear and nonlinear SIC, based on a multi-layered joint channel estimator structure. In the case of linear SI, given that both the independent and dependent variables are subject to measurement errors, and considering the impact of impulsive noise present in wireless environments on the convergence process of adaptive algorithms, the m-MTLS algorithm is proposed to enhance robustness. For the nonlinear distortion generated by the PA, we combine the joint channel estimator and generalized HP basis function to propose a joint SIC method that outperforms current state-of-the-art SIC algorithms in scenarios of both high and low ISR, as well as high SNR. However, the complexity of this method is higher, mainly due to the increase in the number of basis function coefficients and the length of the filter during joint estimation, as well as the number of layers *M*, which also contributes to the increased complexity.

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

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

**Acknowledgments:** The authors are sincerely grateful to the editor and anonymous reviewers for their insightful and constructive comments that helped us improve this work greatly.

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

#### References

- Liu, G.; Yu, F.R.; Ji, H.; Leung, V.C.M.; Li, X. In-Band Full-Duplex Relaying: A Survey, Research Issues and Challenges. *IEEE Commun. Surv. Tutor.* 2015, 17, 500–524. [CrossRef]
- Lin, Z.; Niu, H.; An, K.; Wang, Y.; Zheng, G.; Chatzinotas, S.; Hu, Y. Refracting RIS-Aided Hybrid Satellite-Terrestrial Relay Networks: Joint Beamforming Design and Optimization. *IEEE Trans. Aerosp. Electron. Syst.* 2022, 58, 3717–3724. [CrossRef]
- Ma, R.; Yang, W.; Guan, X.; Lu, X.; Song, Y.; Chen, D. Covert mmWave Communications With Finite Blocklength Against Spatially Random Wardens. *IEEE Internet Things J.* 2024, 11, 3402–3416. [CrossRef]
- Guo, K.; Dong, C.; An, K. NOMA-Based Cognitive Satellite Terrestrial Relay Network: Secrecy Performance Under Channel Estimation Errors and Hardware Impairments. *IEEE Internet Things J.* 2022, *9*, 17334–17347. [CrossRef]
- Sabharwal, A.; Schniter, P.; Guo, D.; Bliss, D.W.; Rangarajan, S.; Wichman, R. In-Band Full-Duplex Wireless: Challenges and Opportunities. *IEEE J. Sel. Areas Commun.* 2014, 32, 1637–1652. [CrossRef]
- Kolodziej, K.E.; Perry, B.T.; Herd, J.S. In-Band Full-Duplex Technology: Techniques and Systems Survey. *IEEE Trans. Microw. Theory Tech.* 2019, 67, 3025–3041. [CrossRef]
- Anttila, L.; Lampu, V.; Hassani, S.A.; Campo, P.P.; Korpi, D.; Turunen, M.; Pollin, S.; Valkama, M. Full-Duplexing with SDR Devices: Algorithms, FPGA Implementation, and Real-Time Results. *IEEE Trans. Wirel. Commun.* 2021, 20, 2205–2220. [CrossRef]
- Laco, J.M.; Gregorio, F.H.; González, G.; Cousseau, J.E.; Riihonen, T.; Wichman, R. Patch Antenna Design for Full-Duplex Transceivers. In Proceedings of the 2017 European Conference on Networks and Communications (EuCNC), Oulu, Finland, 12–15 June 2017; pp. 1–5.
- Le, A.T.; Tran, L.C.; Huang, X.; Guo, Y.J.; Hanzo, L. Analog Least Mean Square Adaptive Filtering for Self-Interference Cancellation in Full Duplex Radios. *IEEE Wirel. Commun.* 2021, 28, 12–18. [CrossRef]
- Song, S.; Tang, Y.; Wei, X.; Zhou, Y.; Lu, X.; Wang, Z.; Ge, S. Digital Self-Interference Cancellation With Robust Multi-layered Total Least Mean Squares Adaptive Filters. In Proceedings of the 2023 IEEE 98th Vehicular Technology Conference (VTC2023-Fall), Hong Kong, China, 10–13 October 2023; pp. 1–5.
- 11. Hongbo, L.; Jin, M.; Songhu, G.; Junkai, M. Identification and Simulation of Interference Cancellation System Self Adaptive Filtering. In Proceedings of the 2018 37th Chinese Control Conference (CCC), Wuhan, China, 25–27 July 2018; pp. 1892–1897.
- Yongjiang, L.; Luhao, B.; Dong, Z. Adaptive Digital Self-Interference Cancellation Based on Fractional Rrder LMS in LFMCW Radar. J. Syst. Eng. Electron. 2021, 32, 573–583. [CrossRef]

- 13. Duarte, M.; Dick, C.; Sabharwal, A. Experiment-Driven Characterization of Full-Duplex Wireless Systems. *IEEE Trans. Wirel. Commun.* **2012**, *11*, 4296–4307. [CrossRef]
- Korpi, D.; Choi, Y.S.; Huusari, T.; Anttila, L.; Talwar, S.; Valkama, M. Adaptive Nonlinear Digital Self-Interference Cancellation for Mobile Inband Full-Duplex Radio: Algorithms and RF Measurements. In Proceedings of the 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 6–10 December 2015; pp. 1–7.
- Lee, H.; Lee, N. Nonlinear and Non-Stationary Self-Interference Cancellation for Full-Duplex Wireless Systems. In Proceedings of the ICC 2023—IEEE International Conference on Communications, Rome, Italy, 28 May–1 June 2023; pp. 1370–1375.
- 16. Bharadia, D.; McMilin, E.; Katti, S. Full duplex radios. SIGCOMM Comput. Commun. Rev. 2013, 43, 375–386. [CrossRef]
- Li, L.; Zhao, H. A Robust Total Least Mean M-Estimate Adaptive Algorithm for Impulsive Noise Suppression. *IEEE Trans. Circuits Syst. II Express Briefs* 2020, 67, 800–804. [CrossRef]
- 18. Ding, L. Digital Predistortion of Power Amplifiers for Wireless Applications; Georgia Institute of Technology: Atlanta, GA, USA, 2004.
- Korpi, D.; Huusari, T.; Choi, Y.S.; Anttila, L.; Talwar, S.; Valkama, M. Digital Self-Interference Cancellation Under Nonideal RF Components: Advanced Algorithms and Measured Performance. In Proceedings of the 2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Stockholm, Sweden, 28 June–1 July 2015; pp. 286–290.
- Korpi, D.; Anttila, L.; Syrjälä, V.; Valkama, M. Widely Linear Digital Self-Interference Cancellation in Direct-Conversion Full-Duplex Transceiver. *IEEE J. Sel. Areas Commun.* 2014, 32, 1674–1687. [CrossRef]
- Komatsu, K.; Miyaji, Y.; Uehara, H. Frequency-Domain Hammerstein Self-Interference Canceller for In-Band Full-Duplex OFDM Systems. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017; pp. 1–6.
- 22. Yimin, H.; Hongzhi, Z.; Shihai, S. Frequency-Domain Nonlinear Interference Cancellation Scheme with IQ Imbalance in Full-Duplex Systems. *Sci. Sin. Inform.* 2023, 53, 1212–1229.
- Towliat, M.; Guo, Z.; Cimini, L.J.; Xia, X.G.; Song, A. Enhanced Accuracy of Self-Interference Cancellation in Underwater Acoustic Full-Duplex Communication. In Proceedings of the MILCOM 2022—2022 IEEE Military Communications Conference (MILCOM), Rockville, MD, USA, 28 November–2 December 2022; pp. 781–786.
- 24. Despina-Stoian, C.; Youssef, R.; Digulescu, A.; Radoi, E.; Gautier, R.; Serbanescu, A. Widely-Linear Digital Self-Interference Cancellation in Full-Duplex USRP Transceiver. *Sensors* **2022**, *22*, 9607. [CrossRef] [PubMed]
- Arablouei, R.; Werner, S.; Doğançay, K. Analysis of the Gradient-Descent Total Least-Squares Adaptive Filtering Algorithm. *IEEE Trans. Signal Process.* 2014, 62, 1256–1264. [CrossRef]
- Shongwe, T.; Vinck, A.J.H.; Ferreira, H.C. A Study on Impulse Noise and Its Models. SAIEE Afr. Res. J. 2015, 106, 119–131. [CrossRef]
- 27. Saleh, A. Frequency-Independent and Frequency-Dependent Nonlinear Models of TWT Amplifiers. *IEEE Trans. Commun.* **1981**, 29, 1715–1720. [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.