SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

November 2022, Vol. 65 212301:1–212301:17
https://doi.org/10.1007/s11432-021-3303-0

Performance analysis of the nonlinear self-interference cancellation for full-duplex communications

Nanzhou HU, Shanghui XIAO, Wensheng PAN, Qiang XU, Shihai SHAO* & Youxi TANG

National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731*, China*

Received 12 April 2021/Revised 16 June 2021/Accepted 20 July 2021/Published online 26 October 2022

Abstract As the impairments of hardware circuits, such as the nonlinearities of power amplifiers (PAs), limit the self-interference suppression performance in full-duplex systems, nonlinear self-interference cancellation (SIC) has attracted much research attention. According to some existing studies, nonlinear SIC in full-duplex systems can be implemented with either nonlinear modeling or radio frequency (RF) signal feedback. However, to the best of our knowledge, there is no theoretical analysis and comparison of the cancellation performance with the two methods. In this paper, the performance of the digital nonlinear SIC with RF signal feedback and nonlinear modeling is analyzed and compared for the first time. The theoretical SIC capabilities of the two methods are derived, and the closed-form solutions are obtained. The factors affecting the performance of the two methods are discussed with the theoretical analysis. Then, by simulations, the theoretical results are verified and the performances of the nonlinear SIC with the two methods are compared in different environments.

Keywords full duplex, self-interference cancellation, PA nonlinearity, digital domain, performance analysis.

Citation Hu N Z, Xiao S H, Pan W S, et al. Performance analysis of the nonlinear self-interference cancellation for full-duplex communications. Sci China Inf Sci, 2022, 65(11): 212301, https://doi.org/10.1007/s11432-021-3303-0

1 Introduction

Driven by the accelerating process of informatization, researches on sixth generation (6G) wireless communication networks have begun, in which enhanced spectral efficiency is an important performance metrics [1]. In recent years, in-band full-duplex technology has received extensive attention and been a focus of research in the wireless communication domain [2–23], as it is capable of improving the spectrum efficiency and flexibility [2, 3]. The self-interference (SI) signal generated by simultaneous transmission and reception at the same frequency, generally composed of linear and nonlinear components, is the main factor hindering full-duplex communication, and effective suppression of the SI is the basis of the implementation of full-duplex technology [2, 3]. A three-stage self-interference cancellation technique is usually adopted in typical duplex systems [3], which involves the cancellation in propagation domain [4,5], analog domain [6, 7] and digital domain [9–22], and high cancellation capabilities without consideration of nonlinear SI have been achieved in existing researches [4–7]. However, in high-power communication systems, the nonlinear SI components, including phase noise, in-phase/quadrature (IQ) imbalance and power amplifier (PA) nonlinearity, are the main factors causing a technical bottleneck for full-duplex SI cancellation (SIC), and the impact of PA nonlinearity is the largest [8]. Research shows that when the transmit power is 25 dBm and a total of 60 dB of cancellation capability is provided in the propagation and radio frequency (RF) domains, the power of the rest of the SI nonlinear components is still approximately 41 dB higher than the noise floor after digital SIC [8]. Therefore, nonlinear SIC has attracted much research attention, and some nonlinear SIC methods have been proposed [9–22], among which the

* Corresponding author (email: ssh@uestc.edu.cn)

SIC with RF signal feedback and that with nonlinear modeling are the two main methods for nonlinear SIC in existing researches.

The SIC with RF signal feedback uses a feedback channel at the end of the RF transmission link, through which the transmission signal including impairments of the RF circuits, such as the nonlinearity in the PA, can be transmitted back to the digital domain as a reference signal and then convolved with the SI channel estimation result so that the SI can be reconstructed and canceled. With the help of the feedback channel, Refs. [9–11] achieved a nonlinear SIC performance of 0–3 dB above the noise level at a transmit power of 20–30 dBm and a bandwidth of 20–100 MHz. The SIC with RF signal feedback does not need to consider the complex characteristics of nonlinearity. However, additional hardware costs are introduced, and some hardware imperfections in the feedback channel will affect the cancellation performance.

The SIC with nonlinear modeling uses the baseband signal as a reference signal, models the nonlinear SI, and estimates the model parameters from the received SI to reconstruct the SI and cancel it in the presence of nonlinearity. Refs. [12–22] used different nonlinear models to achieve a simulated nonlinear SIC performance of 0–4.3 dB above the noise level at a transmit power of 10–30 dBm and a bandwidth of 10–80 MHz. SIC with nonlinear modeling does not require the assistance of RF circuits but relies on appropriate nonlinear modeling and has a high computational complexity. Generally, only odd and low-order components in the nonlinearity are considered, as they have larger power and smaller separation in the spectrum from the intended signal, making them difficult to filter out. However, for broadband signals, high-order nonlinear components are also difficult to remove with a filter. Moreover, as the transmit power increases, the power of high-order components will grow faster than that of low-cost components according to the nonlinearity model [24]. When the order of the nonlinear components higher than the noise floor in the SI is greater than the modeling order, the SIC with nonlinear modeling cannot guarantee sufficient SIC performance.

To further guide engineering practice, investigations into the performance bounds and influencing factors of nonlinear SIC with these two methods are needed. However, to the best of our knowledge, most of the current researches have focused on the explanation and implementation of nonlinear SIC methods and lack a corresponding theoretical analysis and comparison of the cancellation performance. In this paper, we analyze the theoretical performance and influencing factors of nonlinear SIC with RF signal feedback, which is significant as a reference for the analysis and selection of nonlinear SIC methods in engineering. The main contributions are summarized as follows.

- First, we consider the main relevant impairments of RF circuits, including PA nonlinearity, and analyze the performance of nonlinear SIC with RF signal feedback and nonlinear modeling. After derivation in details, the closed-form solutions of the cancellation capabilities with two methods are obtained which have not been proposed in existing researches. It is shown that the cancellation capabilities with the two methods have the same upper bound, but are limited by different factors.

- Then, based on the closed-form expressions of the cancellation capabilities with two methods, we discussed the factors affecting the performance of the two methods for the first time, respectively. We compare the two SIC methods in terms of SIC performance, hardware cost and computational complexity. Some suggestions of which method should be used under different conditions are provided for the implementation of full-duplex engineering for the first time.

- Finally, by simulation, the theoretical results are verified, the performance of nonlinear SIC with the two methods are compared in different environments. The results have certain reference value to engineering practice.

The rest of this paper is organized as follows. In Section 2, the model of SI including nonlinearity is provided. Sections 3 and 4 introduce the principles and analyze the theoretical cancellation performance of the SIC with RF signal feedback and nonlinear modeling, respectively. Section 5 first analyzes the pros and cons of the two methods from a theoretical point of view and then shows the numerical results under different parameter environments. Finally, Section 6 summarizes the full text and gives conclusion.

Notations. The following notations are used in this paper. Matrixes and vectors are denoted as bold capital letters and $\mathbf{I}$ is the identity matrix. $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$, and $(\cdot)^*$ are the transpose, the conjugate transpose, and the complex conjugate, respectively. $\mathrm{E}[\cdot]$ means the statistical expectation and $\otimes$ denotes the convolution. $\mathrm{Tr}[\cdot]$ and $\|\cdot\|$ are the trace and the F norm of a matrix, respectively.

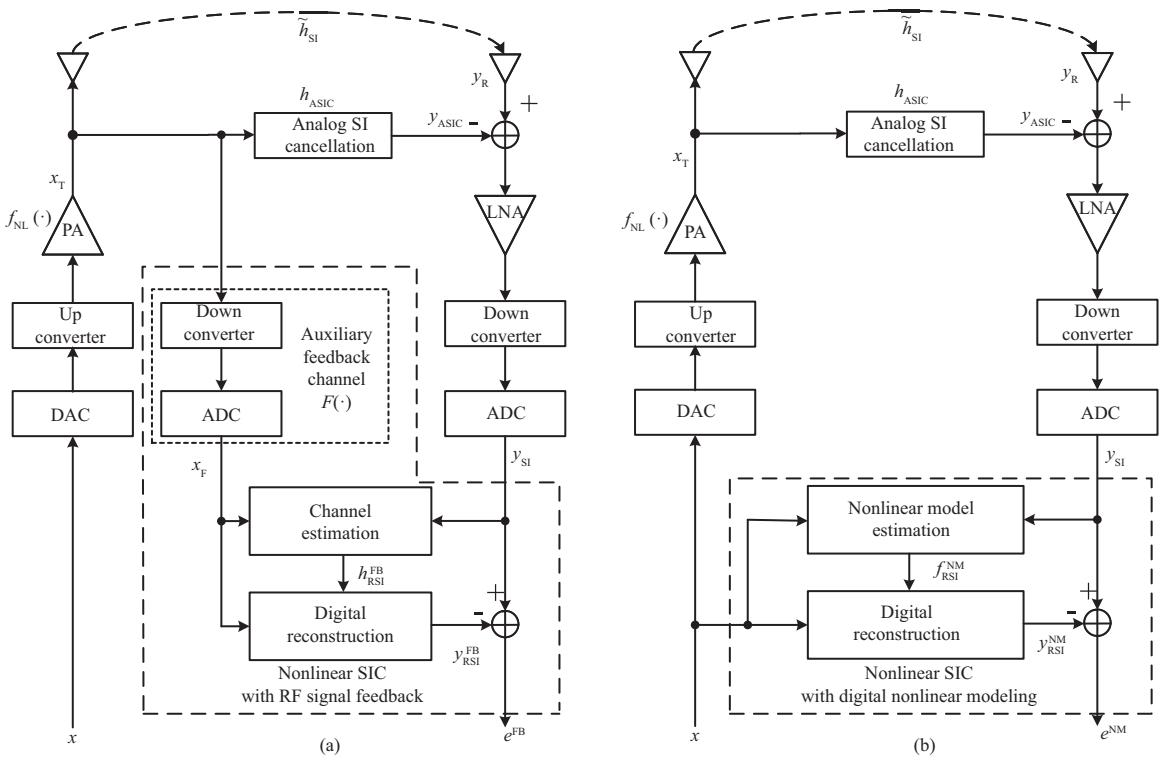

Figure 1 Block diagrams of systems using nonlinear SIC with (a) RF signal feedback and (b) nonlinear modeling

2 SI model and cancellation capability

2.1 SI signal model

Figure 1 depicts the architecture of the full-duplex systems using nonlinear SIC with feedback from an RF signal and nonlinear modeling. As shown in Figure 1(a) and (b), after the baseband signal is upconverted to the RF domain, nonlinear components such as the phase noise, IQ imbalance, and PA nonlinearity are introduced. By denoting the digital baseband signal as x, the local transmission signal that includes nonlinearity at the end of the transmission link can be expressed as

$$x_{\mathrm{T}}(n) = x_{\mathrm{L}}(n) + x_{\mathrm{NL}}(n) + w_{\mathrm{T}}(n), \tag{1}$$

where $x_{\mathrm{L}}(n)$ and $x_{\mathrm{NL}}(n)$ are the linear and nonlinear components, respectively, and $w_{\mathrm{T}}(n)$ is the noise of the transmission link with variance $\sigma^2_{\mathrm{R},w}$.

The combination of linear and nonlinear components can be described by the memory polynomial model [24], which is given by

$$x_{\mathrm{T}}(n) = f_{\mathrm{NL}}[x(n)] + w_{\mathrm{T}}(n) = \sum_{\substack{p=1 \\ p\,\text{is odd}}}^{P_{\mathrm{NL}}} \sum_{m=0}^{M_{\mathrm{NL}}} f_{p,m}\psi_{p,m}[x(n)] + w_{\mathrm{T}}(n), \tag{2}$$

where $f_{\mathrm{NL}}(\cdot)$ is the transfer function of the nonlinear process, P_{NL} is the maximum order of the nonlinearity, M_{NL} is the memory depth, $\psi_{p,m}$ is the coefficient of order p and delay m, and

$$\psi_{p,m}[x(n)] = x(n-m)\,|x(n-m)|^{p-1}. \tag{3}$$

After the local signal is transmitted and propagates through the SI channel, SI is introduced into the receiving antenna. Suppose that the response of the SI channel is $\tilde{h}_{\mathrm{SI}}$; then, the received SI signal is

$$y_{\mathrm{R}}(n) = \tilde{h}_{\mathrm{SI}}(n) \otimes x_{\mathrm{T}}(n) + w_{\mathrm{R}}(n), \tag{4}$$

where w_{R} is the noise of the receiving link and has variance $\sigma^2_{\mathrm{R},w}$.

Since the reference signal is drawn from the end of the transmission link, analog SIC can simultaneously suppress the linear and nonlinear SI components. Suppose that the response of analog SIC is h_{ASIC}; then, the SI signal reconstructed in the RF domain is

$$y_{\mathrm{ASIC}}(n) = h_{\mathrm{ASIC}}(n) \otimes x_{\mathrm{T}}(n). \tag{5}$$

After the analog SIC in the RF domain, quantization noise is introduced into the SI during quantization. Unless otherwise specified, all the SI mentioned in the following text refers to the SI signal received in the digital domain, which is given by

$$\begin{aligned} y_{\mathrm{SI}}(n) &= y_{\mathrm{R}}(n) - y_{\mathrm{ASIC}}(n) + q_{\mathrm{R}}(n) \\ &= [\tilde{h}_{\mathrm{SI}}(n) - h_{\mathrm{ASIC}}(n)] \otimes x_{\mathrm{T}}(n) + w_{\mathrm{R}}(n) + q_{\mathrm{R}}(n) \\ &= h_{\mathrm{SI}}(n) \otimes x_{\mathrm{T}}(n) + w_{\mathrm{R}}(n) + q_{\mathrm{R}}(n), \end{aligned} \tag{6}$$

where

$$h_{\mathrm{SI}}(n) = \tilde{h}_{\mathrm{SI}}(n) - h_{\mathrm{ASIC}}(n) \tag{7}$$

is the response of the SI channel after RF domain SIC, and $q_{\mathrm{R}}(n)$ is the quantization noise. In this paper, the quantization noise is regarded as Gaussian noise independent of the signals, and it is established for uniform quantization and general bandpass signals [25]. Suppose that the variance of the quantization noise is $\sigma^2_{\mathrm{R},q}$, the signal transmit power is P_s, and the total gain of the SI channel after RF cancellation is P_h; then, the interference-to-noise ratio of the receiving channel is given by

$$\eta_{\mathrm{R}} = \frac{P_s P_h}{\sigma^2_{\mathrm{R}}}, \tag{8}$$

where $\sigma^2_{\mathrm{R}} = \sigma^2_{\mathrm{R},w} + \sigma^2_{\mathrm{R},q} + P_h \sigma^2_{\mathrm{T},w}$ and $P_h = G_h / G_{\mathrm{RF}}$ is determined by the SI channel gain G_h, the total SIC capability in the propagation domain and RF domain G_{RF}. The SIC in the propagation domain and RF domain is usually realized by physical isolation, beamforming, SI reconstruction with RF circuit, etc. [3]. Here, we ignore the specific method in the propagation domain and RF domain, and only give a reasonable assumption used in [8] that the total SIC capability in the propagation domain and RF domain $G_{\mathrm{RF}} \approx 60$ dB. Generally, the SI channel gain G_h is small so that $P_h \ll 1$. Thus, $\sigma^2_{\mathrm{R}} = \sigma^2_{\mathrm{R},w} + \sigma^2_{\mathrm{R},q}$ is established, and the influence of the transmission link noise in the received SI signal can be ignored.

2.2 Cancellation capability

In this article, the relative power ratio of the SI signal to the residual SI and noise is used to evaluate the nonlinear SI digital cancellation performance, and the SIC capability is defined as

$$\rho = 10 \lg \frac{P_{\mathrm{SI}}}{P_e}, \tag{9}$$

where P_{SI} is the power of the received SI signal, and P_e is the power of the residual SI and noise after SIC. Then, the performance of the SIC with RF signal feedback and nonlinear modeling is defined as

$$\rho^{\mathrm{FB}} = 10 \lg \frac{P_{\mathrm{SI}}}{P_e^{\mathrm{FB}}}, \tag{10}$$

and

$$\rho^{\mathrm{NM}} = 10 \lg \frac{P_{\mathrm{SI}}}{P_e^{\mathrm{NM}}}, \tag{11}$$

respectively.

The following two sections introduce the principles of the SIC with RF signal feedback and that with nonlinear modeling and analyse their theoretical cancellation performance and influencing factors.

3 SIC with RF signal feedback

The idea of nonlinear SIC with RF signal feedback is to couple the signal at the end of the transmission link back to the digital domain as a reference signal. Since the reference signal contains nonlinear components, the nonlinear SI can be canceled with only channel estimation and linear reconstruction. However, quantization noise and thermal noise are introduced into the reference signal in the feedback channel; thus, the reference signal obtained in the digital domain is not an ideal local transmission signal, which affects the cancellation performance. In this section, the expressions of related signals in the cancellation process are given first, and then, the theoretical cancellation performance is derived and analysed.

3.1 Models of related signals

Suppose that the feedback response is $F(\cdot)$; then, the actual reference signal is $x_{\mathrm{F}}(n)=F[x_{\mathrm{T}}(n)]$, which can be expressed by

$$x_{\mathrm{F}}(n)=F(x_{\mathrm{T}}(n))=x_{\mathrm{T}}(n)+w_{\mathrm{F}}(n)+q_{\mathrm{F}}(n), \tag{12}$$

where $w_{\mathrm{F}}(n)$ is the thermal noise of the feedback channel with variance $\sigma^2_{\mathrm{F},w}$ and $q_{\mathrm{F}}(n)$ is the quantization noise of the feedback channel with variance $\sigma^2_{\mathrm{F},q}$. Then, the signal-to-noise ratio of the feedback channel can be given by

$$\eta_{\mathrm{F}}=\frac{P_s}{\sigma^2_{\mathrm{F}}}, \tag{13}$$

where $\sigma^2_{\mathrm{F}}=\sigma^2_{\mathrm{F},w}+\sigma^2_{\mathrm{F},q}+\sigma^2_{\mathrm{T},w}$.

After the transmitted signal is fed back to the digital domain as a reference, it is used to estimate the SI channel together with the received SI signal and then convolved with the estimated SI channel response to reconstruct the SI, namely,

$$y^{\mathrm{FB}}_{\mathrm{RSI}}(n)=h^{\mathrm{FB}}_{\mathrm{RSI}}(n)\otimes x^{\mathrm{FB}}_{\mathrm{F}}(n), \tag{14}$$

where $h^{\mathrm{FB}}_{\mathrm{RSI}}$ is the channel estimation result (SI reconstruction response function) and $y^{\mathrm{FB}}_{\mathrm{RSI}}(n)$ is the reconstructed SI signal.

Then, the residual SI can be given by

$$\begin{aligned} e^{\mathrm{FB}}(n)&=y_{\mathrm{SI}}(n)-y^{\mathrm{FB}}_{\mathrm{RSI}}(n)\\ &=\left[h_{\mathrm{SI}}(n)\otimes x_{\mathrm{T}}(n)-h^{\mathrm{FB}}_{\mathrm{RSI}}(n)\otimes x_{\mathrm{F}}(n)\right]+w_{\mathrm{R}}(n)+q_{\mathrm{R}}(n)\\ &=\left[h_{\mathrm{SI}}(n)-h^{\mathrm{FB}}_{\mathrm{RSI}}(n)\right]\otimes x_{\mathrm{T}}(n)+z_{\mathrm{R}}(n)-h^{\mathrm{FB}}_{\mathrm{RSI}}(n)\otimes z_{\mathrm{F}}(n), \end{aligned} \tag{15}$$

where

$$z_{\mathrm{R}}(n)=w_{\mathrm{R}}(n)+q_{\mathrm{R}}(n), \tag{16}$$

$$z_{\mathrm{F}}(n)=w_{\mathrm{F}}(n)+q_{\mathrm{F}}(n). \tag{17}$$

3.2 Analysis of the cancellation capability

Eq. (15) shows that the power of the residual SI is determined by the channel estimation error, thermal noise and quantization noise of the receiving and feedback channels. We can write (15) in vector and matrix form with a number of samples N and channel length L:

$$\boldsymbol{e}^{\mathrm{FB}}=\boldsymbol{y}_{\mathrm{SI}}-\boldsymbol{y}^{\mathrm{FB}}_{\mathrm{RSI}}=\boldsymbol{X}_{\mathrm{T}}\left(\boldsymbol{h}_{\mathrm{SI}}-\boldsymbol{h}^{\mathrm{FB}}_{\mathrm{RSI}}\right)-\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}^{\mathrm{FB}}_{\mathrm{RSI}}+\boldsymbol{z}_{\mathrm{R}}, \tag{18}$$

where $\boldsymbol{e}^{\mathrm{FB}}$, $\boldsymbol{y}_{\mathrm{SI}}$, $\boldsymbol{y}^{\mathrm{FB}}_{\mathrm{RSI}}$ and $\boldsymbol{z}_{\mathrm{R}}$ are vectors composed of N samples of e^{FB}, y_{SI}, $y^{\mathrm{FB}}_{\mathrm{RSI}}$ and z_{R}, respectively, $\boldsymbol{h}_{\mathrm{SI}}$ and $\boldsymbol{h}^{\mathrm{FB}}_{\mathrm{RSI}}$ are vectors of SI channel and reconstructed SI channel with the channel length L, respectively,

$$\boldsymbol{X}_{\mathrm{T}}=\begin{bmatrix} x_{\mathrm{T}}(0) & x_{\mathrm{T}}(-1) & \cdots & x_{\mathrm{T}}(-L+1)\\ x_{\mathrm{T}}(1) & x_{\mathrm{T}}(0) & \cdots & x_{\mathrm{T}}(-L+2)\\ \vdots & \vdots & \ddots & \vdots\\ x_{\mathrm{T}}(N-1) & x_{\mathrm{T}}(N-2) & \cdots & x_{\mathrm{T}}(N-L) \end{bmatrix}, \tag{19}$$

$$\boldsymbol{Z}_{\mathrm{F}}=\begin{bmatrix} z_{\mathrm{F}}(0) & z_{\mathrm{F}}(-1) & \cdots & z_{\mathrm{F}}(-L+1) \\ z_{\mathrm{F}}(1) & z_{\mathrm{F}}(0) & \cdots & z_{\mathrm{F}}(-L+2) \\ \vdots & \vdots & \ddots & \vdots \\ z_{\mathrm{F}}(N-1) & z_{\mathrm{F}}(N-2) & \cdots & z_{\mathrm{F}}(N-L) \end{bmatrix}. \tag{20}$$

The least-squares method is adopted for channel estimation, and the result is given by

$$\begin{aligned} \boldsymbol{h}_{\mathrm{RSI}}^{\mathrm{FB}} &= \left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{y}_{\mathrm{SI}}^{0} \\ &= \left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\left(\boldsymbol{X}_{\mathrm{T}}^{0}\boldsymbol{h}_{\mathrm{SI}}^{0}+\boldsymbol{z}_{\mathrm{R}}^{0}\right) \\ &= \boldsymbol{h}_{\mathrm{SI}}^{0}+\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{z}_{\mathrm{R}}^{0}-\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{Z}_{\mathrm{F}}^{0}\boldsymbol{h}_{\mathrm{SI}}^{0}. \end{aligned} \tag{21}$$

In practical communication, channel estimation is performed with the pilot symbols in a frame. In the above equation, the superscript 0 is used to indicate that the vectors and matrix are of pilot symbols. The channel estimation result is used for cancellation in a coherent period, and the SI channel can remain approximately unchanged in a coherent period.

After some derivation and approximation (as shown in Appendix A), the power of the residual SI and noise can be obtained as

$$P_{e}^{\mathrm{FB}}=\mathrm{E}\left\|\boldsymbol{e}^{\mathrm{FB}}\right\|^{2}\approx N\left(\sigma_{\mathrm{R}}^{2}+P_{h}\sigma_{\mathrm{F}}^{2}\right)\left(1+\frac{L}{N-L}\right). \tag{22}$$

We transform the above expression to obtain the relative energy (power) ratio between the SI signal and the residual SI, which can be used to evaluate the nonlinear SI digital cancellation performance. The SIC capability is defined as

$$\begin{aligned} \rho^{\mathrm{FB}} &= 10\lg\frac{P_{\mathrm{SI}}}{P_{e}^{\mathrm{FB}}}\approx 10\lg\frac{\eta_{\mathrm{R}}\eta_{\mathrm{F}}}{(\eta_{\mathrm{R}}+\eta_{\mathrm{F}})\left(1+\frac{L}{N-L}\right)} \\ &= 10\lg\eta_{\mathrm{R}}+10\lg\eta_{\mathrm{F}}-10\lg\left(\eta_{\mathrm{R}}+\eta_{\mathrm{F}}\right)-10\lg\left(1+\frac{L}{N-L}\right), \end{aligned} \tag{23}$$

where $P_{\mathrm{SI}}=P_{s}P_{h}$ is the power of the received SI signal.

The above equation indicates that the cancellation performance is related to the interference-to-noise ratio of the receiving channel, the signal-to-noise ratio of the feedback channel, the number of sampling points used for estimation and the number of paths in the SI channel. The larger the number of channel paths and the smaller the number of sampling points used for estimation, the worse the SIC performance is. Generally, the number of sampling points used in channel estimation is much larger than the number of paths in the SI channel, that is, $N\gg L$; thus,

$$\rho^{\mathrm{FB}}\approx 10\lg\eta_{\mathrm{R}}+10\lg\eta_{\mathrm{F}}-10\lg\left(\eta_{\mathrm{R}}+\eta_{\mathrm{F}}\right). \tag{24}$$

Therefore, in general, the cancellation performance is limited only by the interference-to-noise ratio of the receiving channel η_{R} and the signal-to-noise ratio of the feedback channel η_{F}. Let us consider the extreme cases first to obtain an intuitive conclusion. Two situations arise, as shown below:

$$\rho^{\mathrm{FB}}\approx\begin{cases} 10\lg\eta_{\mathrm{R}},\ \eta_{\mathrm{F}}\gg\eta_{\mathrm{R}}, \\ 10\lg\eta_{\mathrm{F}},\ \eta_{\mathrm{R}}\gg\eta_{\mathrm{F}}. \end{cases} \tag{25}$$

For $\eta_{\mathrm{F}}\gg\eta_{\mathrm{R}}$, the SIC capability is infinitely close to the interference-to-noise ratio of the receiving channel, which means that the SI is almost completely canceled; for $\eta_{\mathrm{R}}\gg\eta_{\mathrm{F}}$, the SIC capability is approximately the interference-to-noise ratio of the feedback channel. Based on the above, the upper bound of the SIC capability is the smaller of η_{F} and η_{R} in the extreme cases, and noise in the feedback channel will cause the SIC capability to be lower than the interference-to-noise ratio of the receiving channel.

Assuming that the same devices are used in the receiving and feedback channels, the noise floors of the two can be regarded as consistent. However, since the intended signal and SI exist in the receiving channel at the same time, the dynamic range cannot be fully used as in the feedback channel during quantization, so the interference-to-noise ratio of the receiving channel will be lower than the signal-to-noise ratio of the feedback channel. Generally, $\eta_{\mathrm{R}} \leqslant \eta_{\mathrm{F}}$ holds. Specifically, the feedback channel is generally higher than the receiving channel by approximately 10 dB or more in engineering practice. Therefore, we have

$$\begin{aligned}\rho^{\mathrm{FB}} &\approx 10\lg\eta_{\mathrm{R}} + 10\lg\eta_{\mathrm{F}} - 10\lg\left(\eta_{\mathrm{R}} + \eta_{\mathrm{F}}\right) = 10\lg\eta_{\mathrm{R}} + 10\lg\frac{1}{\eta_{\mathrm{R}}/\eta_{\mathrm{F}} + 1} \\ &\geqslant 10\lg\eta_{\mathrm{R}} + 10\lg\frac{1}{2} = 10\lg\eta_{\mathrm{R}} - 3 = 10\lg\frac{\eta_{\mathrm{R}}}{2} = 10\lg\frac{P_{\mathrm{SI}}}{2P_{w,\mathrm{R}}},\end{aligned} \tag{26}$$

where $P_{w,\mathrm{R}}$ is the power of the noise of the receiving channel. The above inequality indicates that the SI can be canceled to less than 3 dB above the signal-to-noise ratio of an ideal full-duplex system through SIC with RF signal feedback when $\eta_{\mathrm{R}} \leqslant \eta_{\mathrm{F}}$ holds. The residual SI level is equal to the noise level of the receiving channel when the equation is established; that is, the SI is canceled to the noise floor. Generally, in actual engineering applications, it is enough that the SI is suppressed to below the noise floor.

In conclusion, the range of the capability of SIC with RF signal feedback is

$$10\lg\eta_{\mathrm{R}} - 3 \leqslant \rho^{\mathrm{FB}} \approx 10\lg\eta_{\mathrm{R}} + 10\lg\eta_{\mathrm{F}} - 10\lg\left(\eta_{\mathrm{R}} + \eta_{\mathrm{F}}\right) < 10\lg\eta_{\mathrm{R}}. \tag{27}$$

For $\eta_{\mathrm{R}} = \eta_{\mathrm{F}}$, the equal sign on the left holds, and the power of the residual SI is equal to that of the noise of the receiving channel, with a 3 dB loss of the signal-to-noise ratio at the receiver; as $\eta_{\mathrm{F}} - \eta_{\mathrm{R}}$ increases, the SIC capability ρ tends toward the interference-to-noise ratio η_{R}. For $\eta_{\mathrm{F}} \gg \eta_{\mathrm{R}}$, the SI is almost completely canceled.

4 SIC with digital nonlinear modelling

In nonlinear SIC with nonlinear modeling, the baseband signal is used as the reference signal, i.e., $x_{\mathrm{ref}}^{\mathrm{NM}}(n) = x(n)$, and the memory polynomial model is used to reconstruct the SI including nonlinearity. After the signal is received in the digital domain, the reference signal is used together with the received SI signal to estimate the model coefficients, and the SI is reconstructed with the memory polynomial model. The SIC with nonlinear modeling does not need the assistance of additional RF circuits but requires the model to have a higher fitting degree to the actual nonlinearity; otherwise, nonlinear components that cannot be canceled will exist. In addition, the computational complexity of this method is relatively high. In this section, the expressions of related signals in the cancellation process are given first, and then, the theoretical cancellation performance is derived and analysed.

4.1 Models of related signals

Suppose that the transfer function of the nonlinear SI reconstruction process is $f_{\mathrm{RSI}}^{\mathrm{NM}}(\cdot)$; the reconstruction signal can be expressed by

$$y_{\mathrm{RSI}}^{\mathrm{NM}}(n) = f_{\mathrm{RSI}}^{\mathrm{NM}}\left[x(n)\right] = \sum_{\substack{p=1 \\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{RSI}}^{\mathrm{NM}}} \sum_{m=0}^{M_{\mathrm{RSI}}^{\mathrm{NM}}} f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}} \psi_{p,m}\left[x(n)\right] + w_{\mathrm{T}}(n), \tag{28}$$

where $\psi_{p,m}$ is defined in (3), $f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}}$ is the polynomial coefficient of the nonlinear SI reconstruction process, $P_{\mathrm{RSI}}^{\mathrm{NM}}$ is the maximum polynomial order of the reconstruction, and $M_{\mathrm{RSI}}^{\mathrm{NM}}$ is the reconstruction memory depth.

Then, the residual SI of the final output is

$$\begin{aligned}
&e^{\mathrm{NM}}(n)\\
&=y_{\mathrm{SI}}(n)-y_{\mathrm{RSI}}^{\mathrm{NM}}(n)\\
&=h_{\mathrm{SI}}(n)\otimes x_{\mathrm{PA}}(n)-f_{\mathrm{RSI}}^{\mathrm{NM}}(x(n))+z_{\mathrm{R}}(n)\\
&=\sum_{l=0}^{L}h_{\mathrm{SI}}(l)\sum_{\substack{p=1\\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{NL}}}\sum_{m=0}^{M_{\mathrm{NL}}}f_{p,m}\psi_{p,m}[x(n-l)]-\sum_{\substack{p=1\\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{RSI}}^{\mathrm{NM}}}\sum_{m=0}^{M_{\mathrm{RSI}}^{\mathrm{NM}}}f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}}\psi_{p,m}[x(n)]+z_{\mathrm{R}}(n)\\
&=\sum_{\substack{p=1\\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{NL}}}\sum_{m=0}^{M_{\mathrm{NL}}+L}\left(\sum_{l=0}^{\min\{L,m\}}h_{\mathrm{SI}}(l)f_{p,m-l}\right)\psi_{p,m}[x(n)]-\sum_{\substack{p=1\\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{RSI}}^{\mathrm{NM}}}\sum_{m=0}^{M_{\mathrm{RSI}}^{\mathrm{NM}}}f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}}\psi_{p,m}[x(n)]+z_{\mathrm{R}}(n),
\end{aligned}\tag{29}$$

where

$$z_{\mathrm{R}}(n)=w_{\mathrm{R}}(n)+q_{\mathrm{R}}(n).\tag{30}$$

According to the above formula, the joint response of the nonlinear process and the SI channel can be expressed by a memory polynomial model with the same basis functions, same order and increased memory depth. The coefficients become the convolution of the original memory polynomial coefficients and the SI channel coefficients. Therefore, the entire SI transfer process can be directly modeled, estimated and reconstructed with memory polynomials. However, as the transmit power increases, the high-order nonlinear power that was originally lower than the noise floor will gradually increase until it affects the receiving end, which may cause the nonlinear modeling order in practical applications to be less than the actual order, resulting in the inability to completely cancel the SI. Since the memory depth of PAs is often limited, the memory depth considered in the SIC with nonlinear modeling can be assumed to be sufficient, that is, $M_{\mathrm{RSI}}\geqslant M_{\mathrm{NL}}+L$. For the convenience of expression, the follow-up is performed after satisfying this assumption, and the order and memory depth of the SIC with nonlinear modeling are recorded as P and $M+L$. Then,

$$e^{\mathrm{NM}}(n)=\sum_{\substack{p=1\\ p\,\mathrm{is\,odd}}}^{P}\sum_{m=0}^{M+L}\left[\left(\sum_{l=0}^{\min\{L,m\}}h_{\mathrm{SI}}(l)f_{p,m-l}\right)-f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}}\right]\psi_{p,m}(x(n))+z_{\mathrm{HO}}(n)+w_{\mathrm{R}}(n)+q_{\mathrm{R}}(n),\tag{31}$$

where z_{HO} is the high-order nonlinearity that cannot be canceled due to insufficient modeling order, which can be expanded as

$$z_{\mathrm{HO}}(n)=\sum_{\substack{p=P+2\\ p\,\mathrm{is\,odd}}}^{P_{\mathrm{NL}}}\sum_{m=0}^{M+L}\left(\sum_{l=0}^{\min\{L,m\}}h_{\mathrm{SI}}(l)f_{p,m-l}\right)\psi_{p,m}(x(n)).\tag{32}$$

4.2 Analysis of the cancellation capability

Eq. (31) shows that the residual SI of the digital modeling method depends on the reconstruction order P, actual order P_{NL}, reconstruction memory depth M, actual memory depth M_{NL}, and estimation error of the polynomial coefficient $f_{\mathrm{RSI}\,p,m}^{\mathrm{NM}}$. It can be written as the vector or matrix form of the corresponding signal formed by N sampling points, that is,

$$\boldsymbol{e}=\boldsymbol{y}_{\mathrm{SI}}-\boldsymbol{y}_{\mathrm{RSI}}^{\mathrm{NM}}=\boldsymbol{\Psi}\left(\boldsymbol{f}-\boldsymbol{f}_{\mathrm{RSI}}^{\mathrm{NM}}\right)+\boldsymbol{z}_{\mathrm{R}}+\boldsymbol{z}_{\mathrm{HO}},\tag{33}$$

where

$$\begin{aligned}
\boldsymbol{\Psi}&=\begin{bmatrix}\psi_{1,0}(x(0)) & \psi_{3,0}(x(0)) & \cdots & \psi_{P,M+L}(x(0))\\ \psi_{1,0}(x(1)) & \psi_{3,0}(x(1)) & \cdots & \psi_{P,M+L}(x(1))\\ \vdots & \vdots & \ddots & \vdots\\ \psi_{1,0}(x(N-1)) & \psi_{3,0}(x(N-1)) & \cdots & \psi_{P,M+L}(x(N-1))\end{bmatrix}\\
&=\begin{bmatrix}x(0) & x(0)|x(0)|^2 & \cdots & x(-M)|x(-M-L)|^{P-1}\\ x(1) & x(1)|x(1)|^2 & \cdots & x(1-M)|x(1-M-L)|^{P-1}\\ \vdots & \vdots & \ddots & \vdots\\ x(N-1) & x(N-1)|x(N-1)|^2 & \cdots & x(N-1-M)|x(N-1-M-L)|^{P-1}\end{bmatrix}
\end{aligned}\tag{34}$$

is the memory polynomial matrix of the baseband signal,

$$\boldsymbol{f} = \left[f_{1,0} \ f_{1,1} \ \cdots \ f_{P,M} \right]^{\mathrm{T}} \tag{35}$$

and

$$\boldsymbol{f}_{\mathrm{RSI}}^{\mathrm{NM}} = \left[f_{\mathrm{RSI}\,1,0}^{\mathrm{NM}} \ f_{\mathrm{RSI}\,1,1}^{\mathrm{NM}} \ \cdots \ f_{\mathrm{RSI}\,P,M+L}^{\mathrm{NM}} \right]^{\mathrm{T}} \tag{36}$$

are the memory polynomial coefficients obtained during the actual transmission and the estimation (the model of the convolution of the PA response and the channel response), and

$$\boldsymbol{z}_{\mathrm{HO}} = \left[z_{\mathrm{HO}}(0) \ z_{\mathrm{HO}}(1) \ \cdots \ z_{\mathrm{HO}}(N-1) \right]^{\mathrm{T}}. \tag{37}$$

Same as Subsection 3.2, the subscript 0 indicates that the vectors and matrix are of pilot symbols, and the least-squares estimation is used; then,

$$\begin{aligned} \boldsymbol{f}_{\mathrm{RSI}}^{\mathrm{NM}} &= \left[(\boldsymbol{\Psi}^0)^{\mathrm{H}} \boldsymbol{\Psi}^0 \right]^{-1} (\boldsymbol{\Psi}^0)^{\mathrm{H}} \boldsymbol{Y}_{\mathrm{SI}}^0 \\ &= \left[(\boldsymbol{\Psi}^0)^{\mathrm{H}} \boldsymbol{\Psi}^0 \right]^{-1} (\boldsymbol{\Psi}^0)^{\mathrm{H}} \left(\boldsymbol{\Psi}^0 \boldsymbol{f} + \boldsymbol{z}_{\mathrm{R}}^0 \right). \end{aligned} \tag{38}$$

With some derivation and approximation in Appendix B, the power of the residual SI and noise can be expressed as

$$\begin{aligned} P_{\mathrm{RSI}}^{\mathrm{NM}} &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{M+L+1}{N} \sigma_{\mathrm{R}}^2 \mathrm{Tr} \left[\boldsymbol{R}_V^0 \cdot (\boldsymbol{R}_V)^{-1} \right] \\ &\approx \left[1 + \frac{(M+L+1)(P+1)}{2N} \right] \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}}. \end{aligned} \tag{39}$$

The power ratio of the SI and residual SI ρ is used to evaluate the nonlinear SIC performance, namely,

$$\begin{aligned} \rho^{\mathrm{NM}} &= 10 \lg \frac{E_{\mathrm{SI}}}{P_{\mathrm{RSI}}^{\mathrm{NM}}} \\ &= 10 \lg \frac{N P_h P_s}{[N + (M+L+1)(P+1)/2N] \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}}} \\ &= 10 \lg \eta_{\mathrm{R}} + 10 \lg \eta_{\mathrm{HO}} - 10 \lg \left[\left(1 + \frac{(M+L+1)(P+1)}{2N} \right) \eta_{\mathrm{R}} + \eta_{\mathrm{HO}} \right], \end{aligned} \tag{40}$$

where $\eta_{\mathrm{HO}} = \frac{P_{\mathrm{SI}}}{P_{\mathrm{HO}}}$ is the power ratio between the total SI and residual high-order nonlinear SI components. The above formula shows that the cancellation performance of the SIC with nonlinear modeling is related to the interference-to-noise ratio of the receiving channel, the power ratio between the total SI and residual high-order nonlinear SI components, the number of sampling points used for estimation, the number of channel paths, and the order and memory depth of the nonlinear modeling. Generally, the number of sampling points used for estimation can be sufficiently large, that is, $N \gg (M+L+1)(P+1)/2$; then,

$$\begin{aligned} \rho^{\mathrm{NM}} &= 10 \lg \eta_{\mathrm{R}} + 10 \lg \eta_{\mathrm{HO}} - 10 \lg \left[\left(1 + \frac{(M+L+1)(P+1)}{2N} \right) \eta_{\mathrm{R}} + \eta_{\mathrm{HO}} \right] \\ &\approx 10 \lg \eta_{\mathrm{R}} + 10 \lg \eta_{\mathrm{HO}} - 10 \lg (\eta_{\mathrm{R}} + \eta_{\mathrm{HO}}). \end{aligned} \tag{41}$$

The cancellation performance of the SIC with nonlinear modeling is limited by the interference-to-noise ratio of the receiving channel and the ratio of the SI to residual high-order nonlinear SI components. The residual high-order nonlinearities mainly depend on the transmit power and the modeling order. As the transmit power increases, the power of high-order nonlinearities gradually increases, and when this power approaches or exceeds the noise floor, the performance of the receiver will deteriorate. Therefore, the nonlinear order of which nonlinear components actually affect the performance of the receiver increases with the transmit power. If the modeling order was greater than the maximum order of actual influential nonlinearities, then no residual high-order nonlinearities would exist. Therefore, the cancellation capability of the SIC with nonlinear modeling will be

$$\rho^{\mathrm{NM}} = 10 \lg \eta_{\mathrm{R}}; \tag{42}$$

Table 1 Comparison of the SIC with nonlinear modeling and the SIC with RF signal feedback

	SIC with nonlinear modeling	SIC with RF signal feedback
Upper-bound of SIC capability	Interference-to-noise ratio	Interference-to-noise ratio
Affected factors	High transmit power and deficiency of the modeling order	Hardware impairments in the feedback channel and 0∼3 dB signal-to-noise ratio loss for general devices
Computational complexity	$\mathrm{O}\left[\left(P+1\right)\left(M+L+1\right)/2\right]^3\right]$	$\mathrm{O}\left(L^3\right)$
Hardware cost	No need for additional hardware	A complete feedback channel

that is, it will be equal to the interference-to-noise ratio of the receiving channel, which can be regarded as the SI being canceled completely. However, if the modeling order was less than the maximum order of actual influential nonlinearities, then residual high-order nonlinearities would exist that could not be effectively canceled, thereby reducing the nonlinear SIC capability. Due to the lack of understanding of the statistical characteristics of high-order nonlinearities with changing transmit power, specific mathematical expressions regarding this issue are not presented in this article and will observe it in subsequent simulations.

5 Comparison and discussion

Sections 3 and 4 introduce the two main nonlinear SIC methods and derive their theoretical cancellation capabilities. In this section, the performances of the two methods are compared based on the above derivation results, and the analysis results are verified through simulation. The two methods will be compared from four aspects, which are shown in Table 1.

5.1 Theoretical analysis and comparison

Eqs. (24) and (41) give the theoretical performance expressions of the two nonlinear cancellation methods. Comparing the two, the SIC capabilities of the two nonlinear SIC methods are similar in the structure of the theoretical expressions but are limited by the feedback channel noise and residual high-order nonlinearity, respectively.

The SIC with nonlinear modeling is not affected by the additional noise in the feedback channel. However, the order of the nonlinear components that actually deteriorate the performance of the receiver increases with increasing transmit power, and if the modeling order was less than the actual order, then the performance of the receiver would be affected because the nonlinear components with higher order could not be effectively canceled by the SIC with nonlinear modeling. Therefore, to ensure an effective SIC, the SIC with nonlinear modeling should increase the modeling order as the transmit power increases. When the modeling order is sufficiently large, the SI can be completely canceled.

The SIC with RF signal feedback directly returns the local transmitted signal containing nonlinearity to the digital domain; thus, it is not influenced by the specific nonlinear characteristics, but the feedback process will introduce some noise. If the signal-to-noise ratio of the feedback channel was close to or even smaller than the interference-to-noise ratio of the receiving channel, then the noise introduced in the feedback channel would result in a decrease in the SIC capability. The SIC with RF signal feedback requires a larger dynamic range and quantization of the devices of the feedback channel. Generally, in engineering, the feedback channel has a higher signal-to-noise ratio than the receiving channel, which is sufficient to ensure that SI is canceled close to the noise floor with a 0∼3 dB signal-to-noise ratio loss.

In addition to the SIC performance, the SIC with nonlinear modeling has a higher computational complexity, which is $\mathrm{O}[[(P+1)(M+L+1)/2]^3]$, while that of the SIC with RF signal feedback is only $\mathrm{O}(L^3)$. According to the above analysis, high-transmit-power communication systems often require a higher modeling order P, which leads to a further increase in the computational complexity.

Hardware cost is also a consideration in engineering practice. Compared with the SIC with nonlinear modeling, the SIC with RF signal feedback will introduce additional hardware costs, especially in FD-MIMO communication systems. In practical engineering, there is an independent power amplifier on each antenna branch. Therefore, each antenna branch needs a feedback channel to get the different PA nonlinearities in each branch back to digital domain. It means that the additional hardware costs are proportional to the number of antennas for the SIC with RF signal feedback.

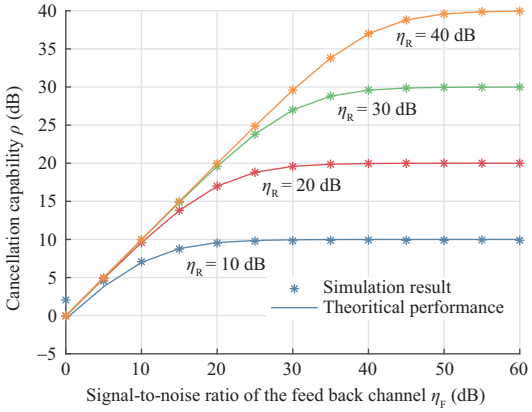

Figure 2 Capability ρ vs. signal-to-noise ratio η_{F} with transmit power = 25 dBm.

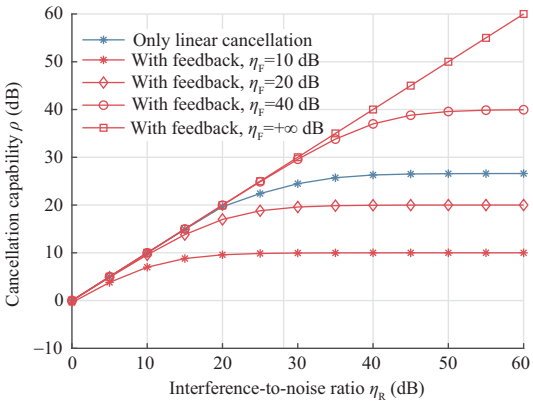

Figure 3 Capability ρ with RF signal feedback and only a linear process vs. interference-to-noise ratio η_{R}.

As higher computational complexity is more likely to accept than more hardware costs, we think the SIC with nonlinear modeling is a better choice for engineering in most cases. However, if the FD radio works under a rapidly time-varying environment, real-time parameter estimation is necessary to obtain enough SIC capability. For this scenario, the SIC with RF signal feedback has superiority because lower computational complexity means less parameter estimation delay.

5.2 Simulation results and discussion

In this section, the capability of SIC with RF signal feedback and only a linear process is simulated with MATLAB under different parameters. In the simulations, the transmitted signal is modulated by quadrature phase shift keying (QPSK), with a bandwidth of $B = 200$ MHz and a carrier frequency of fc = 20 GHz. The PA has a gain of 25 dB and an input third-order intercept point (IIP3) of 15 dBm, and the memory polynomial model is used for modeling and for simulation of the PA with a maximum order of $P = 9$ and a memory depth of $M = 5$. The model TDL-D in the 3Rd Generation Partnership Project Technical Report (3GPP TR) 38.901 is used as the SI channel, and the amounts of SI suppression in the propagation and RF domains are set to 30 and 20 dB, respectively, which are easy to implement with existing research [26]. The characteristics of the SI channel and the PA are assumed to be stable in an estimation period. The nonlinear SI digital cancellation capability is evaluated by the power ratio of the SI and residual SI, which is defined in (9).

Figure 2 shows the SIC capability ρ with respect to η_{F} under different η_{R} values. The results indicate that the theoretical results of the cancellation performance are in good agreement with the simulation results when η_{F} is greater than 10 dB. However, when η_{F} is smaller, some deviation is observed between the simulation results and the theoretical curve. This deviation is due to the assumption that η_{F} is sufficiently large, which is used in the derivation of the theoretical performance and is reasonable for engineering applications.

In addition, as shown in Figure 2, the SIC capability exhibits a logarithmic increase with η_{F}, which conforms to the mathematical relationship shown in (24). For $\eta_{\mathrm{F}} < \eta_{\mathrm{R}}$, the low signal-to-noise ratio of the feedback channel is the main reason for the poor SIC performance, and the SIC performance increases almost linearly with increasing η_{F}. For $\eta_{\mathrm{F}} = \eta_{\mathrm{R}}$, the cancellation performance is approximately 3 dB lower than η_{R}, which means that the SI is suppressed to the noise floor; for $\eta_{\mathrm{F}} > \eta_{\mathrm{R}}$, the SIC capability continues to approach the interference-to-noise ratio, ensuring almost complete SIC. In addition, for $\eta_{\mathrm{F}} \geqslant \eta_{\mathrm{R}} + 10\,\mathrm{dB}$, the difference between the SIC capability ρ and the interference-to-noise ratio η_{R} can be largely ignored, which means that the condition for complete SIC can be relaxed in actual engineering implementation.

In Figure 3, the capability of SIC with RF signal feedback under different η_{F} values and that of SIC with only a linear process versus η_{R} are compared. SIC with only a linear process does not use the feedback channel for assistance but uses the baseband signal as a reference signal and performs SIC through only linear operations. In the simulation, the transmit power is set to 30 dBm, and the nonlinear component (approximately 9.81 dBm) is approximately 20 dB lower than the linear component

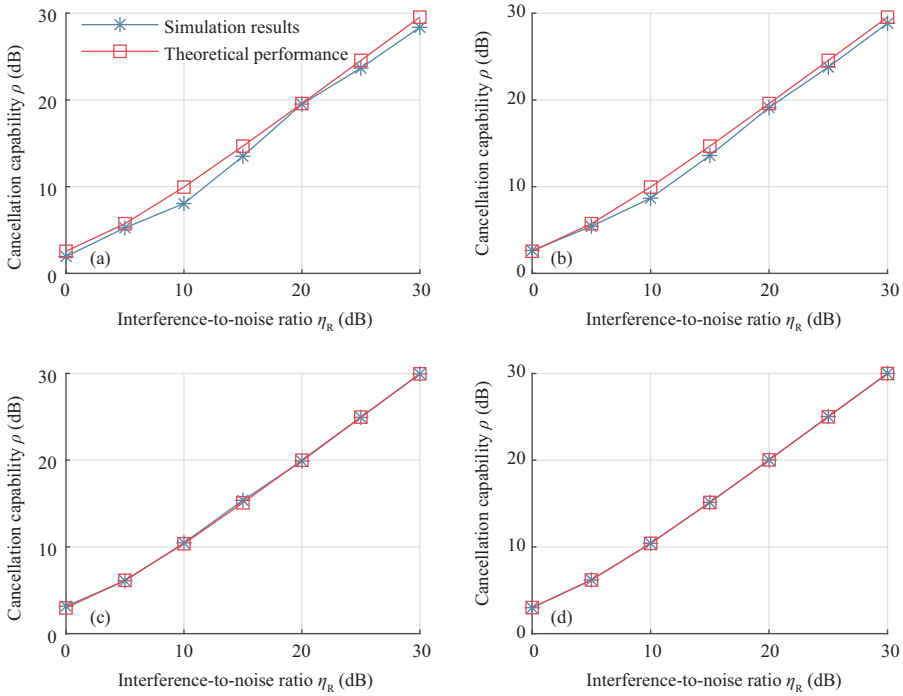

Figure 4 (Color online) Capability ρ with nonlinear modeling vs. interference-to-noise ratio η_{R} with deferent sampling points. (a) $N = 160$; (b) $N = 800$; (c) $N = 1600$; (d) $N = 16000$.

(approximately 29.95 dBm). When η_{F} is less than 20 dB, the noise of the receiving channel is greater than the nonlinear component of the SI, and the nonlinear component may hardly affect the operation of the receiver. Therefore, the capability of SIC with only a linear process is close to that of SIC with RF signal feedback for $\eta_{\mathrm{F}} = 40, +\infty$ dB and is even higher than that of SIC with RF signal feedback for $\eta_{\mathrm{F}} = 10, 20$ dB. When η_{R} is greater than 20 dB, the nonlinear component starts to degrade the performance of the receiver, and the capability of SIC with only a linear process is lower than that of SIC with RF signal feedback for $\eta_{\mathrm{F}} = 40, +\infty$ dB. The simulation results show that the upper bound of the linear cancellation performance is approximately 26.6 dB, which is greater than the power ratio of the linear SI component to the nonlinear SI component of 20 dB, indicating that linear cancellation can cancel some parts of the nonlinear component, but its cancellation capability is limited. When the signal-to-noise ratio of the feedback channel is much greater than the power ratio of the linear component to the nonlinear component of the SI, SIC with RF signal feedback can be guaranteed to be better than SIC with only a linear process.

Figure 4 simulates the performance of the SIC with nonlinear modeling with changing interference-to-noise ratio of the receiving channel under the condition that the transmit power is 25 dBm, the nonlinear modeling order is $P = 9$, and the memory depth is $M = 5$. The results show that under the set parameters, the performance of the SIC with nonlinear modeling based on the memory polynomial model increases with increasing signal-to-noise ratio (interference-to-noise ratio) of the receiving channel, and the SI can be canceled to the noise floor within the range of 0–30 dB. In addition, when $N = 160$, the simulation result is significantly lower than the theoretical performance, up to 2 dB, which is due to the error caused by the smaller number of sample points for estimation, consistent with the analysis in Section 4. With an increasing number of sampling points for estimation, the degree of fitting between the simulation results and theoretical performance gradually improves. For $N = 1600$, the fit to the theoretical performance has basically been achieved. In actual engineering, the number of sampling points required for estimation is not difficult to achieve.

In the scenario of a receiving channel interference-to-noise ratio of 30 dB and three different PAs, Figure 5 compares the cancellation performances of the SIC with RF signal feedback under different signal-to-noise ratios of the feedback channel, the SIC with nonlinear modeling under different modeling orders, and the only linear SIC. For the same transmit power, when the gain and IIP3 of the PA are smaller, the power of nonlinear SI components is greater, and the nonlinear SI components in the three

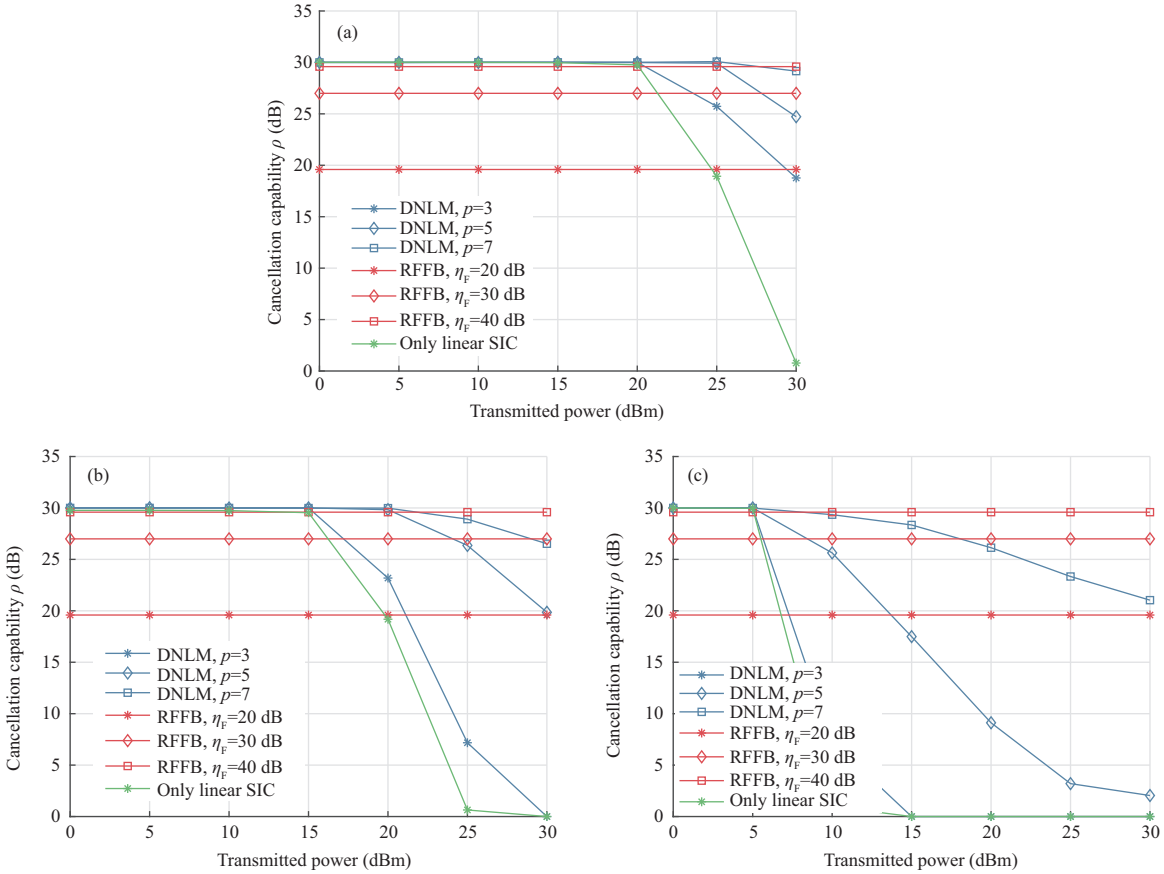

Figure 5 Capability ρ of the two nonlinear SIC methods and only linear SIC vs. transmit power with different PAs. (a) PA gain = 25 dB, IIP3 = 15 dBm; (b) PA gain = 25 dB, IIP3 = 10 dBm; (c) PA gain = 15 dB, IIP3 = 10 dBm.

scenarios shown in Figure 5 above gradually increase in sequence. Figure 5 shows that in the three scenarios shown in (a)–(c), when the transmit power is greater than 20, 15, and 5 dBm, respectively, the capability of the only linear SIC begins to drop significantly, as the nonlinear components begin to be higher than the receiver noise floor. In addition, the capability of the only linear SIC is close to 0 dB when the transmit power is 30, 25, and 15 dBm in the three scenarios because the nonlinear components with higher power are difficult to cancel and affect the channel estimation accuracy. This result shows that for a full-duplex communication system with high transmit power, only linear SIC can hardly address the SI effectively, and nonlinear SIC is necessary to ensure normal communication.

In the three scenarios in Figure 5, the SIC with nonlinear modeling has the best cancellation performance at low transmit power, which is almost equal to the interference-to-noise ratio in the receiving channel of 30 dB. However, as the transmit power increases, the high-order nonlinearities under the noise floor at low transmit power become larger and in turn deteriorate the performance of the receiver, and an insufficient modeling order will lead to a decrease in the cancellation performance. The performance of the SIC with nonlinear modeling of $P = 3$ has the fastest decline. In scenarios (a)–(c), when the transmit power is greater than 20, 15, and 5 dBm, the performance begins to decrease. In (a), when the transmit power is 30 dBm, the performance is still 18.8 dB, and in (b) and (c), when the transmit power is greater than 30 and 15 dBm, the cancellation capability is almost 0 dB. The capability of the SIC with nonlinear modeling of $P = 5$ begins to decrease when the transmit power is greater than 25, 20, and 5 dBm in scenarios (a)–(c), respectively. When the transmit power is 30 dBm, the capability is approximately 19.9, 24.7, and 2.1 dB, respectively. The SIC with nonlinear modeling of $P = 7$ in (a) shows effective cancellation capability in the range of 0–30 dBm transmit power. When the transmit power is 30 dBm, the performance reaches a minimum of 29.1 dB. In scenarios (b) and (c), when the transmit power is greater than 25 and 10 dBm, the performance begins to significantly decrease, and when the transmit power is 30 dBm, the performance drops to approximately 24.7 and 21.3 dB, respectively. For the situation of a PA with high transmit power, low gain and low IIP3, the SIC with nonlinear modeling needs a higher

modeling order to ensure effective cancellation capability. The SIC with RF signal feedback is almost unaffected by the changes in the transmit power and PA parameters. The capability remains stable in the range of 0–30 dBm with the three PAs. It is only limited by the signal-to-noise ratio of the feedback channel. When the ratio is 20, 30, and 40 dB, the cancellation performance is approximately 19.6, 27.0, and 29.6 dB, respectively.

For high-power full-duplex communication systems, the digital modeling cancellation method requires a higher modeling order to achieve effective cancellation, but when the order is sufficient, its cancellation performance is optimal, which can be regarded as complete cancellation. The RF feedback SIC method is only limited by the noise of the feedback channel. When the feedback channel is in good condition, the performance is stable, and an increase in power will not decrease the performance. However, its extreme performance requires the feedback channel signal-to-noise ratio to be extremely high, which makes achieving complete SIC difficult. The only linear cancellation method has an excellent linear SIC ability when low transmit power and nonlinearity have not begun to cause the signal-to-noise ratio to decrease, but as the transmit power increases, the nonlinear component power in the SI increases, and only linear cancellation can no longer meet the needs of SIC.

6 Conclusion

This paper mainly analyzes the theoretical cancellation performances of nonlinear SIC with RF signal feedback and nonlinear modeling and simulates the nonlinear SIC under different environmental parameters. The theoretical derivation shows that the performances of the above two methods have the same upper bound. Sufficient sampling points are needed for parameter estimation to ensure excellent SIC performance close to the upper bound, especially for the nonlinear modeling SIC method which has more parameters to estimate. In addition, the performance of the RF signal feedback SIC method is greatly affected by the noise of the feedback channel, while that of the nonlinear modeling SIC method is affected by the power of the nonlinearities with a higher order than the modeling. Moreover, the calculation complexity of the nonlinear modeling SIC method is relatively high while the RF signal feedback SIC method will introduce additional hardware costs, especially in FD-MIMO communication systems. As higher computational complexity is more likely to accept than more hardware costs, we think the SIC with nonlinear modeling is a better choice in engineering. By simulations, the theoretical cancellation performances are verified. The results show that the RF signal feedback SIC method has better performance than the nonlinear modeling SIC method with a low modeling order under the situation of high transmit power and low IIP3 of PA; conversely, the nonlinear modeling cancellation method performance is better since the noise of the feedback channel limits the performance of the nonlinear SIC with RF signal feedback.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61771107, 61701075, 61601064) and National Key R&D Program of China (Grant No. 2018YFB1801903).

References

1 You X, Wang C X, Huang J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts. Sci China Inf Sci, 2021, 64: 110301
2 Xia X, Xu K, Wang Y, et al. A 5G-enabling technology: benefits, feasibility, and limitations of in-band full-duplex mMIMO. IEEE Veh Technol Mag, 2018, 13: 81–90
3 Sabharwal A, Schniter P, Guo D, et al. In-band full-duplex wireless: challenges and opportunities. IEEE J Sel Areas Commun, 2014, 32: 1637–1652
4 Xia X, Xu K, Zhang D, et al. Beam-domain full-duplex massive MIMO: realizing co-time co-frequency uplink and downlink transmission in the cellular system. IEEE Trans Veh Technol, 2017, 66: 8845–8862
5 Xu K, Shen Z, Wang Y, et al. Hybrid time-switching and power splitting SWIPT for full-duplex massive MIMO systems: a beam-domain approach. IEEE Trans Veh Technol, 2018, 67: 7257–7274
6 Amjad M S, Nawaz H, Ozsoy K, et al. A low-complexity full-duplex radio implementation with a single antenna. IEEE Trans Veh Technol, 2018, 67: 2206–2218
7 Kolodziej K E, McMichael J G, Perry B T. Multitap RF canceller for in-band full-duplex wireless communications. IEEE Trans Wireless Commun, 2016, 15: 4321–4334
8 Korpi D, Riihonen T, Syrjala V, et al. Full-duplex transceiver system calculations: analysis of ADC and linearity challenges. IEEE Trans Wireless Commun, 2014, 13: 3821–3836
9 Ahmed E, Eltawil A M. All-digital self-interference cancellation technique for full-duplex systems. IEEE Trans Wireless Commun, 2015, 14: 3519–3532
10 Doane J P, Kolodziej K E, Perry B T. Simultaneous transmit and receive with digital phased arrays. In: Proceedings of IEEE International Symposium on Phased Array Systems and Technology, Waltham, 2016. 1–6
11 Liu Y, Ma W, Quan X, et al. An architecture for capturing the nonlinear distortion of analog self-interference cancellers in full-duplex radios. IEEE Microw Wireless Compon Lett, 2017, 27: 845–847

12 Korpi D, Heino M, Icheln C, et al. Compact inband full-duplex relays with beyond 100 dB self-interference suppression: enabling techniques and field measurements. IEEE Trans Antennas Propagat, 2017, 65: 960–965
13 Anttila L, Korpi D, Syrjala V, et al. Cancellation of power-amplifier induced nonlinear self-interference in full-duplex transceivers. In: Proceedings of Asilomar Conference on Signals, Systems and Computers, Pacific Grove, 2013. 1193–1198
14 Anttila L, Korpi D, Antonio-Rodriguez E, et al. Modeling and efficient cancellation of nonlinear self-interference in MIMO full-duplex transceivers. In: Proceedings of IEEE Globecom Workshops, Austin, 2014. 777–783
15 Emara M, Roth K, Baltar L G, et al. Nonlinear digital self-interference cancellation with reduced complexity for full duplex systems. In: Proceedings of the 21st International ITG Workshop on Smart Antennas, Berlin, 2017. 1–6
16 Ahmed E A, Eltawil A M, Sabharwal A. Self-interference cancellation with nonlinear distortion suppression for full-duplex systems. In: Proceedings of Asilomar Conference on Signals, Systems and Computers, Pacific Grove, 2013. 1199–1203
17 Quan X, Liu Y, Chen D, et al. Blind nonlinear self-interference cancellation for wireless full-duplex transceivers. IEEE Access, 2018, 6: 37725–37737
18 Sim M S, Chung M K, Kim D, et al. Nonlinear self-interference cancellation for full-duplex radios: from link-level and system-level performance perspectives. IEEE Commun Mag, 2017, 55: 158–167
19 Shen L, Henson B, Zakharov Y, et al. Adaptive nonlinear equalizer for full-duplex underwater acoustic systems. IEEE Access, 2020, 8: 108169
20 Komatsu K, Miyaji Y, Uehara H. Iterative nonlinear self-interference cancellation for in-band full-duplex wireless communications under mixer imbalance and amplifier nonlinearity. IEEE Trans Wireless Commun, 2020, 19: 4424–4438
21 Yang H, Zhang H, Zhang J, et al. Digital self-interference cancellation based on blind source separation and spectral efficiency analysis for the full-duplex communication systems. IEEE Access, 2018, 6: 43946–43955
22 Vogt H, Enzner G, Sezgin A. State-space adaptive nonlinear self-interference cancellation for full-duplex communication. IEEE Trans Signal Process, 2019, 67: 2810–2825
23 Tong J P, Zhong C J. Full-duplex two-way AF relaying systems with imperfect interference cancellation in Nakagami-m fading channels. Sci China Inf Sci, 2021, 64: 182310
24 Ku H, Kenney J S. Behavioral modeling of nonlinear RF power amplifiers considering memory effects. IEEE Trans Microwave Theor Techn, 2003, 51: 2495–2504
25 Bennett W R. Spectra of quantized signals. Bell Syst Technical J, 1948, 27: 446–472
26 Everett E, Sahai A, Sabharwal A. Passive self-interference suppression for full-duplex infrastructure nodes. IEEE Trans Wireless Commun, 2014, 13: 680–694

Appendix A

In appendix A, the power of the residual SI and noise of SIC with RF signal feedback is derived. As $\boldsymbol{z}_{\mathrm{R}}$, $\boldsymbol{z}_{\mathrm{R}}^{0}$ and $\boldsymbol{Z}_{\mathrm{F}}$ are composed of thermal noise and quantization noise, they can be regarded as completely independent and as following a Gaussian distribution. Therefore, with (18) and (21), the power of the residual SI and noise can be given by

$$\begin{aligned} P_e^{\mathrm{FB}} = \mathrm{E}\left\|\boldsymbol{e}^{\mathrm{FB}}\right\|^2 = \mathrm{E}\left[\mathrm{Tr}\left(\boldsymbol{e}^{\mathrm{FB}}\boldsymbol{e}^{\mathrm{FB\,H}}\right)\right] = N\sigma_{\mathrm{R}}^2 + NP_h\sigma_{\mathrm{F}}^2 + \left(\sigma_{\mathrm{R}}^2 + P_h\sigma_{\mathrm{F}}^2\right)\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\right]\right]\right] \\ - 2\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{Z}_{\mathrm{F}}^{0}\boldsymbol{h}_{\mathrm{SI}}^{0}\left(\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\right)^{\mathrm{H}}\right]\right]\right]. \end{aligned} \tag{A1}$$

To further simplify (A1), by expanding the fourth term of the above formula, the below equation can be obtained from the matrix inversion lemma.

$$\begin{aligned} &\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\left(\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\right)^{\mathrm{H}}\right]\right]\right] \\ &= \mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}\boldsymbol{X}_{\mathrm{F}}^{0\,-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\boldsymbol{h}_{\mathrm{SI}}^{0}\left(\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\right)^{\mathrm{H}}\right]\right]\right] \\ &= \mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\left(\boldsymbol{X}_{\mathrm{T}}+\boldsymbol{Z}_{\mathrm{F}}\right)\left(\boldsymbol{X}_{\mathrm{T}}^{0}+\boldsymbol{Z}_{\mathrm{F}}^{0}\right)^{-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\left(\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\right)^{\mathrm{H}}\right]\right]\right] \\ &= \mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[P_h\sigma_{\mathrm{F}}^2\boldsymbol{X}_{PA}^{0\,-1}\left(\boldsymbol{Z}_{\mathrm{F}}^{0\,-1}+\boldsymbol{X}_{\mathrm{T}}^{0\,-1}\right)^{-1}\boldsymbol{X}_{\mathrm{T}}^{0\,-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\right]\right]\right]. \end{aligned} \tag{A2}$$

The condition number can be used to measure the error of the matrix inversion. When the matrix has a small perturbation δA, the relative deviation between the matrix inversion result and the real inverse matrix is as follows:

$$\frac{\left\|\boldsymbol{A}^{-1}-(\boldsymbol{A}+\boldsymbol{\delta A})^{-1}\right\|}{\|\boldsymbol{A}^{-1}\|} \leqslant \frac{\mathrm{cond}\,(\boldsymbol{A})\,\frac{\|\boldsymbol{\delta A}\|}{\|\boldsymbol{A}\|}}{1-\mathrm{cond}\,(\boldsymbol{A})\,\frac{\|\boldsymbol{\delta A}\|}{\|\boldsymbol{A}\|}}. \tag{A3}$$

From the above formula, the inversion error can be regarded as small when the condition number is close to 1 and the perturbation $\boldsymbol{\delta}A$ is small relative to the matrix $\boldsymbol{A}$. Therefore, when adding a small perturbation to the matrix $\boldsymbol{Z}_{\mathrm{F}}^{0\,-1}$, which is a Gaussian covariance matrix with a condition number of 1, the error of the inversion result can also be considered small. When η_{F} is large, the elements in $\boldsymbol{X}_{\mathrm{T}}^{0\,-1}$ are much smaller than the elements in $\boldsymbol{Z}_{\mathrm{F}}^{0\,-1}$ and $\boldsymbol{Z}_{\mathrm{F}}^{0}$, so $(\boldsymbol{Z}_{\mathrm{F}}^{0\,-1}+\boldsymbol{X}_{\mathrm{T}}^{0\,-1})^{-1} \approx \boldsymbol{Z}_{\mathrm{F}}^{0}$ is established. Therefore,

$$\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}\boldsymbol{X}_{\mathrm{F}}^{0\,-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\boldsymbol{h}_{\mathrm{SI}}^{0}\left(\boldsymbol{Z}_{\mathrm{F}}\boldsymbol{h}_{\mathrm{SI}}^{0}\right)^{\mathrm{H}}\right]\right]\right] \approx \mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[P_h\sigma_{\mathrm{F}}^2\boldsymbol{X}_{\mathrm{T}}^{0\,-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\boldsymbol{X}_{\mathrm{T}}^{0\,-1}\boldsymbol{Z}_{\mathrm{F}}^{0}\right]\right]\right] \approx 0. \tag{A4}$$

With the above approximations, when the return signal noise is relatively large, the following approximation can be considered valid.

$$P_e^{\mathrm{FB}} = \mathrm{E}\left\|\boldsymbol{e}^{\mathrm{FB}}\right\|^2 \approx N\left(\sigma^2+P_h\sigma_{\mathrm{F}}^2\right)+\left(\sigma^2+P_h\sigma_{\mathrm{F}}^2\right)\cdot\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}^{0\,-H}\boldsymbol{X}_{\mathrm{F}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}\boldsymbol{X}_{\mathrm{F}}^{0\,-1}\right]\right]\right], \tag{A5}$$

where $\boldsymbol{X}_{\mathrm{F}}$ follows a Gaussian distribution with mean 0 and variance σ_{F}^2. Thus, $\boldsymbol{X}_{\mathrm{F}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}$ follows a noncentral Wishart distribution with N degrees of freedom, mean $\boldsymbol{X}_{\mathrm{T}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{T}}$ and covariance matrix $P_s+\sigma_{\mathrm{F}}^2\boldsymbol{I}_{\mathrm{L}}$, and $(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0})^{-1}$ follows a noncentral inverse Wishart distribution. Since the mathematical expectation of a noncentral Wishart inverse distribution has no general exact solution,

we assume here that the signal for the channel estimation follows a Gaussian distribution, that is, $\boldsymbol{X}_{\mathrm{T}}^{0} \sim N(0, P_s)$, and thus, $\boldsymbol{X}_{\mathrm{F}}^{0} \sim N(0, P_s+\sigma_{\mathrm{F}}^2)$. Therefore, $(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0})^{-1}$ follows a central inverse Wishart distribution with N degrees of freedom and covariance matrix $(P_s+\sigma_{\mathrm{F}}^2)\boldsymbol{I}_{\mathrm{L}}$, the mathematical expectation of which is given in the paper shown in the footnote[1)]. As $\boldsymbol{X}_{\mathrm{F}}$ and $\boldsymbol{X}_{\mathrm{F}}^{0}$ are signals with different periods, the two can be assumed to be independent; thus,

$$\mathrm{E}\left[\mathrm{Tr}\left[\mathrm{Re}\left[\boldsymbol{X}_{\mathrm{F}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}\left(\boldsymbol{X}_{\mathrm{F}}^{0\,\mathrm{H}}\boldsymbol{X}_{\mathrm{F}}^{0}\right)^{-1}\right]\right]\right]=\mathrm{Tr}\left[\mathrm{Re}\left[\left(N\sigma_{\mathrm{F}}^2\boldsymbol{I}+\boldsymbol{X}_{\mathrm{T}}{}^{\mathrm{H}}\boldsymbol{X}_{\mathrm{T}}\right)\frac{1}{(N-L)\left(P_s+\sigma_{\mathrm{F}}^2\right)}\boldsymbol{I}_{\mathrm{L}}\right]\right]=\frac{NL}{N-L}. \tag{A6}$$

Substituting (A6) into (A5), the residual SI and noise's power can be obtained:

$$P_e^{\mathrm{FB}}=\mathrm{E}\left\|\boldsymbol{e}^{\mathrm{FB}}\right\|^2\approx N\left(\sigma_{\mathrm{R}}^2+P_h\sigma_{\mathrm{F}}^2\right)\left(1+\frac{L}{N-L}\right). \tag{A7}$$

Appendix B

In Appendix B, the power of the residual SI and noise of SIC with nonlinear modeling is derived. According to (38), the coefficient estimation error is only caused by noise, which is recorded as

$$\begin{aligned}\Delta\boldsymbol{f}&=\boldsymbol{f}-\boldsymbol{f}_{\mathrm{RSI}}^{\mathrm{NM}}\\&=\left[(\boldsymbol{\Psi}^0)^{\mathrm{H}}\boldsymbol{\Psi}^0\right]^{-1}(\boldsymbol{\Psi}^0)^{\mathrm{H}}\boldsymbol{z}_{\mathrm{R}}^0.\end{aligned} \tag{B1}$$

Then, the cancellation error can be written as

$$\begin{aligned}\boldsymbol{e}^{\mathrm{NM}}&=\boldsymbol{y}_{\mathrm{SI}}-\boldsymbol{y}_{\mathrm{RSI}}^{\mathrm{NM}}\\&=\boldsymbol{\Psi}f-\boldsymbol{\Psi}\boldsymbol{f}_{\mathrm{RSI}}^{\mathrm{NM}}+\boldsymbol{z}_{\mathrm{R}}\\&=\boldsymbol{\Psi}\Delta\boldsymbol{f}+\boldsymbol{z}_{\mathrm{R}}+\boldsymbol{z}_{\mathrm{HO}}.\end{aligned} \tag{B2}$$

Its energy can be calculated by the following formula:

$$\begin{aligned}P_{\mathrm{RSI}}^{\mathrm{NM}}&=\frac{1}{N}\mathrm{E}\left\|\boldsymbol{e}^{\mathrm{NM}}\right\|^2=\frac{1}{N}\mathrm{E}\left[\mathrm{Tr}\left[\boldsymbol{e}^{\mathrm{NM}}(\boldsymbol{e}^{\mathrm{NM}})^{\mathrm{H}}\right]\right]\\&=\frac{1}{N}\mathrm{E}\Big[\mathrm{Tr}\Big(\boldsymbol{z}_{\mathrm{HO}}\boldsymbol{z}_{\mathrm{HO}}{}^{\mathrm{H}}+\boldsymbol{z}_R\boldsymbol{z}_R^{\mathrm{H}}+\boldsymbol{\Psi}\Delta\boldsymbol{f}\boldsymbol{z}_R{}^{\mathrm{H}}+\boldsymbol{z}_R(\boldsymbol{\Psi}\Delta\boldsymbol{f})^{\mathrm{H}}+\boldsymbol{\Psi}\Delta\boldsymbol{f}(\boldsymbol{\Psi}\Delta\boldsymbol{f})^{\mathrm{H}}\Big)\Big]\\&=\frac{1}{N}\mathrm{Tr}\Big[\mathrm{E}\left[\boldsymbol{z}_{\mathrm{HO}}\boldsymbol{z}_{\mathrm{HO}}{}^{\mathrm{H}}\right]+\mathrm{E}\left[\boldsymbol{z}_R\boldsymbol{z}_R{}^{\mathrm{H}}\right]+2\mathrm{E}\left[\mathrm{Re}\left[(\boldsymbol{\Psi}\Delta\boldsymbol{f})^{\mathrm{H}}\right]\boldsymbol{z}_R\right]+\mathrm{E}\left[\boldsymbol{\Psi}\Delta\boldsymbol{f}(\boldsymbol{\Psi}\Delta\boldsymbol{f})^{\mathrm{H}}\right]\Big]\\&=\sigma_{\mathrm{R}}^2+P_{\mathrm{HO}}+\frac{1}{N}\mathrm{Tr}\left[\mathrm{E}\Big[\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi}\Delta\boldsymbol{f}(\Delta\boldsymbol{f})^{\mathrm{H}}\Big]\right],\end{aligned} \tag{B3}$$

where P_{HO} is the power of residual high-order nonlinearities.

The relevant results are estimated by least squares, and the error covariance matrix is $\mathrm{E}[\Delta\boldsymbol{f}\Delta\boldsymbol{f}^{\mathrm{H}}]=\sigma_{\mathrm{R}}^2[(\boldsymbol{\Psi}^0)^{\mathrm{H}}\boldsymbol{\Psi}^0]^{-1}$. Then, the residual SI is

$$\begin{aligned}P_{\mathrm{RSI}}^{\mathrm{NM}}&=\frac{1}{N}\mathrm{E}\left\|\boldsymbol{e}^{\mathrm{NM}}\right\|^2\\&=\sigma_{\mathrm{R}}^2+P_{\mathrm{HO}}+\frac{\sigma^2}{N}\mathrm{Tr}\left[\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi}\left((\boldsymbol{\Psi}^0)^{\mathrm{H}}\boldsymbol{\Psi}^0\right)^{-1}\right].\end{aligned} \tag{B4}$$

To further consider the magnitude of the residual SI without affecting the calculation results, the polynomial matrix is written as

$$\boldsymbol{\Psi}=\left[\begin{array}{cccc}\boldsymbol{V}^0 & \boldsymbol{V}_1 & \cdots & \boldsymbol{V}_{M+L}\end{array}\right], \tag{B5}$$

where

$$\begin{aligned}\boldsymbol{V}_m&=\begin{bmatrix}\psi_{1,m}(x(0)) & \psi_{3,m}(x(0)) & \cdots & \psi_{P,m}(x(0))\\ \psi_{1,m}(x(1)) & \psi_{3,m}(x(1)) & \cdots & \psi_{P,m}(x(1))\\ \vdots & \vdots & \ddots & \vdots\\ \psi_{1,m}(x(N-1)) & \psi_{3,m}(x(N-1)) & \cdots & \psi_{P,m}(x(N-1))\end{bmatrix}\\&=\begin{bmatrix}x(-m) & & & \\ & x(1-m) & & \\ & & \ddots & \\ & & & x(N-1-m)\end{bmatrix}\cdot\begin{bmatrix}1 & |x(-m)|^2 & \cdots & |x(-m)|^{P-1}\\ 1 & |x(1-m)|^2 & \cdots & |x(1-m)|^{P-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & |x(N-1-m)|^2 & \cdots & |x(N-1-m)|^{P-1}\end{bmatrix},\ m=0,1,\ldots,M\end{aligned} \tag{B6}$$

is the matrix formed by the sub-terms of the signal whose memory depth is m, and

$$\boldsymbol{V}_m{}^{\mathrm{H}}\boldsymbol{V}_m=\tilde{V}_m^{\mathrm{H}}\boldsymbol{\Lambda}\tilde{V}_m=\sum_{n=0}^{N-1}\begin{bmatrix}|x(n-m)|^2 & |x(n-m)|^4 & \cdots & |x(n-m)|^{P+1}\\ |x(n-m)|^4 & |x(n-m)|^6 & \cdots & |x(n-m)|^{P+3}\\ \vdots & \vdots & \ddots & \vdots\\ |x(n-m)|^{P+1} & |x(n-m)|^{P+3} & \cdots & |x(n-m)|^{2P}\end{bmatrix}. \tag{B7}$$

1) Maiwald D, Kraus D. Calculation of moments of complex Wishart and complex inverse Wishart distributed matrices. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Munich, 1997. 3817–3820.

When N is sufficiently large, the above matrix elements can be approximated to the moments of the baseband signal x. Suppose that the signal is a stationary process; then, the value of $\boldsymbol{V}_m{}^{\mathrm{H}}\boldsymbol{V}_m$ is unrelated to m and can be approximated as

$$\boldsymbol{V}_m{}^{\mathrm{H}}\boldsymbol{V}_m \approx N\begin{bmatrix} (x^2) & (x^4) & \cdots & (x^{P+1}) \\ (x^4) & (x^6) & \cdots & (x^{P+3}) \\ \vdots & \vdots & \ddots & \vdots \\ (x^{P+1}) & (x^{P+3}) & \cdots & (x^{2P}) \end{bmatrix} = \boldsymbol{R}_V, \quad \forall m = 0, 1, \ldots, M. \tag{B8}$$

Thus, the polynomial matrix can be expanded into

$$\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi} = \begin{bmatrix} \boldsymbol{v}_0{}^{\mathrm{H}}\boldsymbol{v}_0 & \boldsymbol{v}_0{}^{\mathrm{H}}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_0{}^{\mathrm{H}}\boldsymbol{v}_{M+L} \\ \boldsymbol{v}_1{}^{\mathrm{H}}\boldsymbol{v}_0 & \boldsymbol{v}_1{}^{\mathrm{H}}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_1{}^{\mathrm{H}}\boldsymbol{v}_{M+L} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{v}_{M+L}{}^{\mathrm{H}}\boldsymbol{v}_0 & \boldsymbol{v}_{M+L}{}^{\mathrm{H}}\boldsymbol{v}_1 & \cdots & \boldsymbol{v}_{M+L}{}^{\mathrm{H}}\boldsymbol{v}_{M+L} \end{bmatrix} = \boldsymbol{R}_V \otimes \boldsymbol{R}_{\mathrm{pulse}}, \tag{B9}$$

where $\boldsymbol{R}_{\mathrm{pulse}}$ is the autocorrelation matrix of the pulse waveform, expanded as follows:

$$\boldsymbol{R}_{\mathrm{pulse}} = \begin{bmatrix} R_{\mathrm{pulse}}(0) & R_{\mathrm{pulse}}(-1) & \cdots & R_{\mathrm{pulse}}(-M-L) \\ R_{\mathrm{pulse}}(1) & R_{\mathrm{pulse}}(0) & \cdots & R_{\mathrm{pulse}}(1-M-L) \\ \vdots & \vdots & \ddots & \vdots \\ R_{\mathrm{pulse}}(M+L) & R_{\mathrm{pulse}}(M+L-1) & \cdots & R_{\mathrm{pulse}}(0) \end{bmatrix}, \tag{B10}$$

where R_{pulse} is the autocorrelation function of the pulse waveform, and the matrix size is determined by the memory depth $M+L$. For common pulse waveforms, the matrix is deterministic and invertible.

Substituting the above formula into (B4), we have

$$\begin{aligned} P_{\mathrm{RSI}}^{\mathrm{NM}} &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{\sigma_{\mathrm{R}}^2}{N}\mathrm{Tr}\left[\boldsymbol{\Psi}^{\mathrm{H}}\boldsymbol{\Psi}\left(\boldsymbol{\Psi}^{0\mathrm{H}}\boldsymbol{\Psi}^{0}\right)^{-1}\right] \\ &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{\sigma_{\mathrm{R}}^2}{N}\mathrm{Tr}\left[\boldsymbol{R}_V^0 \otimes \boldsymbol{R}_{\mathrm{pulse}}(\boldsymbol{R}_V \otimes \boldsymbol{R}_{\mathrm{pulse}})^{-1}\right] \\ &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{\sigma_{\mathrm{R}}^2}{N}\mathrm{Tr}\left[\boldsymbol{R}_V^0(\boldsymbol{R}_V)^{-1} \otimes \left(\boldsymbol{R}_{\mathrm{pulse}}\boldsymbol{R}_{\mathrm{pulse}}{}^{-1}\right)\right] \\ &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{M+L+1}{N}\sigma_{\mathrm{R}}^2\mathrm{Tr}\left[\boldsymbol{R}_V^0(\boldsymbol{R}_V)^{-1}\right]. \end{aligned} \tag{B11}$$

Generally, the statistical characteristics of the baseband signal x are unchanged during the estimation period and the cancellation period; thus,

$$\boldsymbol{R}_V^0 \approx \boldsymbol{R}_V \approx N\begin{bmatrix} (x^2) & (x^4) & \cdots & (x^{P+1}) \\ (x^4) & (x^6) & \cdots & (x^{P+3}) \\ \vdots & \vdots & \ddots & \vdots \\ (x^{P+1}) & (x^{P+3}) & \cdots & (x^{2P}) \end{bmatrix}, \quad \forall m = 0, 1, \ldots, M+L. \tag{B12}$$

Thus, the theoretical expression of the residual SI and noise's power is

$$\begin{aligned} P_{\mathrm{RSI}}^{\mathrm{NM}} &= \sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}} + \frac{M+L+1}{N}\sigma_{\mathrm{R}}^2\mathrm{Tr}\left[\boldsymbol{R}_V^0 \cdot (\boldsymbol{R}_V)^{-1}\right] \\ &\approx \left[1 + \frac{(M+L+1)(P+1)}{2N}\right]\sigma_{\mathrm{R}}^2 + P_{\mathrm{HO}}. \end{aligned} \tag{B13}$$