

Air-Induced Passive Intermodulation in FDD MIMO Systems: Algorithms and Measurements

Vesa Lampu¹, Graduate Student Member, IEEE, Lauri Anttila¹, Matias Turunen¹, Marko Fleischer, Jan Hellmann, and Mikko Valkama¹, Fellow, IEEE

Abstract—In this article, we model and develop effective cancellation schemes for passive intermodulation (PIM) distortion induced by external objects in the vicinity of the transmitter—referred to as air-induced PIM—in frequency-division duplex (FDD) multiple-input–multiple-output (MIMO) systems. PIM is interference generated from the transmitted signals undergoing nonlinear transformation, which, in FDD systems, may cause desensitization of the receiver chain. First, we present a general model of the received air-induced PIM signal, with an arbitrary number of dual-carrier TX chains active and an arbitrary number of PIM sources causing interference, on all the intermodulation frequencies. Then, the derived model is used to develop a cancellation scheme based on a complete set of basis functions (BFs) in a rank-2 dual-carrier MIMO system with four active carriers. To alleviate the high complexity of the aforementioned scheme, we then propose a novel alternative cancellation scheme with much reduced complexity, leveraging on the physical modeling of the system, which is capable of handling any number of PIM sources in the system. RF measurement-based experimentations carried out with real-life equipment evidence excellent cancellation capabilities of the complete BF model, which can be retained with much reduced complexity with the proposed alternative technique.

Index Terms—Carrier aggregation (CA), 5G, frequency-division duplex (FDD), multiple-input multiple-output (MIMO), nonlinear distortion cancellation, passive intermodulation (PIM), simultaneous transmit and receive.

I. INTRODUCTION

THE current 5G new radio (NR) specifications support the carrier aggregation (CA) technology, which was introduced already in long-term evolution (LTE)-advanced [1], [2]. CA increases the capacity and link throughput via the use of multiple component carriers (CCs), which can be allocated to frequency bands contiguously or noncontiguously. Allocating the CCs noncontiguously can lead to acute intermodulation distortion when the CCs experience a shared nonlinearity [3],

Manuscript received 8 July 2022; revised 23 September 2022; accepted 13 October 2022. This work was supported in part by the Academy of Finland under Grant 319994, Grant 323461, and Grant 338224; and in part by Nokia Mobile Networks. This article is an extended version from the 2022 IEEE MTT-S International Microwave Symposium (IMS-2022) [17] [DOI: 10.1109/LMWC.2022.3164718]. (Corresponding author: Vesa Lampu.)

Vesa Lampu, Lauri Anttila, Matias Turunen, and Mikko Valkama are with the Department of Electrical Engineering, Tampere University, 33720 Tampere, Finland (e-mail: vesa.lampu@tuni.fi).

Marko Fleischer and Jan Hellmann are with Nokia Mobile Networks, 89081 Ulm, Germany.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMTT.2022.3219459>.

Digital Object Identifier 10.1109/TMTT.2022.3219459

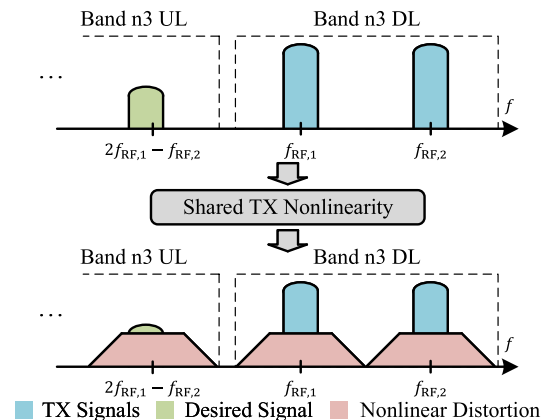


Fig. 1. Concept of intermodulation distortion generation in FDD systems operating at 5G NR band n3. The noncontiguous TX signals propagate through a shared nonlinearity, which generates intermodulation products on the RX band, possibly drowning the actual desired RX signals.

[4]. If the source of the nonlinearity is a passive device, the distortion is called passive intermodulation (PIM) [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16], [17]. Typically, loose connections of cables, filters, switches, antennas [12], [13], [14], [15], [18], [19], [20], [21], or even external metallic objects in the radiation field of the antennas [16], [17] can be the cause of PIM. The PIM caused by external sources is in focus in this work, and such distortion is referred to as air-induced PIM.

In the lower end of the available spectrum, most frequency range 1 (FR1) bands utilize frequency-division duplex (FDD) as the duplexing scheme, as is the case, for example, with 5G NR band n3 [1], [2]. In order to avoid self-interference (SI) directly from the transmitter (TX) signals, the transmission and reception are carried out on different center frequencies in FDD systems. However, such a system is now possibly vulnerable to interference from PIM distortion, as depicted in Fig. 1. If the receiver (RX) operates on or near the intermodulation frequencies, the PIM products can fall on the RX passband, which can possibly desensitize the RX path by drowning the desired received signals from user equipment (UE). All FDD transceivers incorporate a duplexer between the TX and RX chains, as depicted in Fig. 2, which protects the RX path from intermodulation stemming from the TX path. However, any PIM generated after the duplexer (e.g., in connectors, antennas, or nearby metallic objects) can still desensitize the RX path.

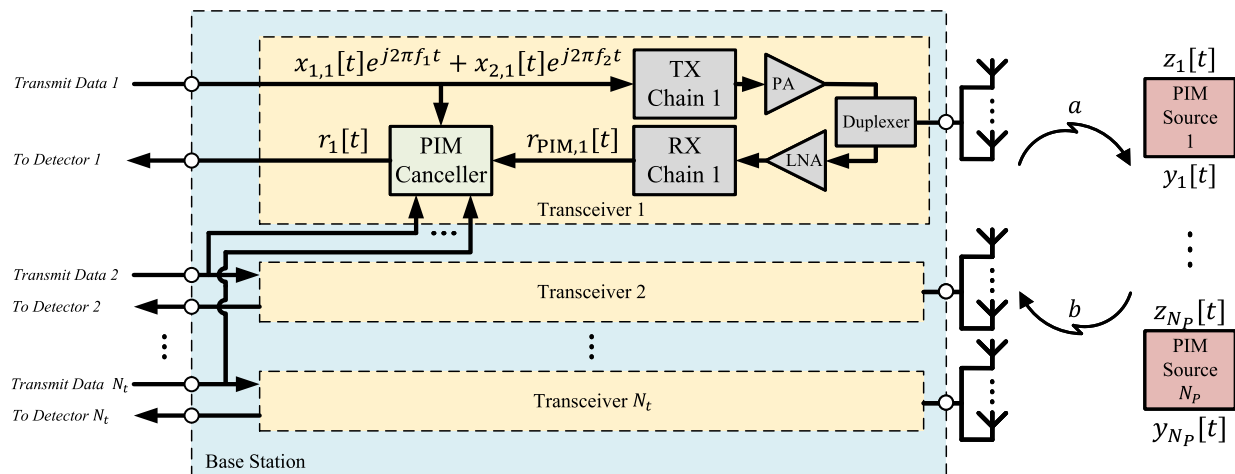


Fig. 2. Simplified block diagram presentation of the considered system, where N_t transceivers located in a BS simultaneously transmit two aggregated CCs, per transceiver, and receive nonlinear interference stemming from N_P air-PIM sources. Some relevant modeling notations are also shown. The PIM canceller block can be implemented, as described in Section III.

Pragmatic suppression of the PIM distortion can be achieved by backing off the TX power; however, such will greatly deteriorate the system power efficiency and the corresponding network coverage and is, therefore, not a congenial solution. Since PIM is commonly generated by the internal circuitry after the duplexer of the transceiver, the simplest solution to such *conducted-PIM* mitigation is to use components that are highly linear. This unfortunately increases the manufacturing costs of the devices, and it is, therefore, not an appealing option [5], [7], [9], [10]. In the *air-induced PIM* scenarios, the obvious solution would be to manually remove the PIM source from the radiation field of the TX, which, in some cases, could turn out to be impossible, if, for example, the PIM source is part of the built environment. Thus, PIM mitigation techniques have been studied in the literature, and they continue to hold scientific and practical relevance in the 5G era. These are shortly reviewed next.

A. Prior Art

In general, the PIM cancellation methods can be divided into analog and digital domain suppressors. In the analog domain, a preliminary work in [5] introduced a nonlinear imposer network, which was shown to suppress PIM by up to 40 dB, while, in [6], an engineered PIM source was used for mitigating PIM. Furthermore, the work in [22] proposed a framework for the adaptation of the modeling and consecutive suppression of high-order PIM distortion. In [7], simulations were carried out to validate PIM mitigation methods by introducing a compensation signal or an additional passive PIM source accompanied by a phase shifter. The works in [23] and [24], in turn, seek to suppress PIM by microstrip coating. Recently, Chen et al. [8] proposed a tunable nonlinear transmission line strip to cancel PIM.

In the digital domain, in [9] and [10], the conducted PIM of two CC systems was modeled with a simple third-order polynomial model and subsequently canceled. The work in [11] modeled and canceled the second-order PIM products of a duplexer. The potential coexistence of nonlinear power amplifiers (PAs) and PIM distortion is considered in [12] and

[18], where reduced complexity models through decoupling of the nonlinearity and memory were introduced. In [16], basic modeling and digital cancellation of air-induced PIM were pursued in a simple two-CC scenario. Finally, our preliminary work in [17] addresses the air-induced PIM cancellation in a multiple-input–multiple-output (MIMO) scenario with four simultaneous carriers.

In addition to PIM modeling and cancellation, the works in [3], [4], [25], [26], [27], [28], [29], and [30] have considered the cancellation of the intermodulation distortion imposed by active components, primarily the TX PA systems. However, these works do not consider PIM, yet it is noted that similar behavioral modeling principles apply when the source of the intermodulation distortion is an active component. Intermodulation distortion can also be mitigated through digital predistortion (DPD), where the inverse of the nonlinear block is applied to the signal digitally, thus linearizing the TX output [31], [32]. In addition, it is mentioned here that inband full-duplex (IBFD) technology [33], [34], [35], where the RX operates simultaneously and on the same frequency as the TX, is dependent on the mitigation of the so-called SI signal, which, without mitigation, desensitizes the RX. In IBFD transceivers, the SI signal is present even if the entire system is completely linear. Therefore, in the IBFD case, one must model the RX signal at the TX frequency, while, in FDD PIM cancellation, the signals are modeled at the intermodulation frequencies.

B. Contributions and Novelty

In this article, building on our initial work in [17], we focus on the modeling and digital cancellation of *air-induced PIM* stemming from the interaction of the radiated electromagnetic wave with external materials outside the transceiver system. Specifically, we consider a challenging rank-2 dual-carrier MIMO FDD system where four CCs are transmitted simultaneously. Besides our initial work in [17], the authors are not aware of any works explicitly aiming to model and cancel air-induced PIM in MIMO scenarios although the issue of PIM has been mentioned in MIMO literature, e.g., in [36],

[37], and [38]. To this end, the so-called full basis function (BF) canceller solution is first described. Then, to alleviate the involved complexity, we also propose an alternative, novel yet simple cancellation method, based on the physical modeling of the system. This method is first described for one dominating air-PIM source, introduced in [17], while it is then also extended to cover multiple simultaneous sources of air-induced PIM. All the models are validated through RF measurements at 5G NR band n3 with true base-station (BS) hardware, evidencing excellent PIM suppression capabilities.

Compared to our initial work in [17] and the overall available literature in the field, the main novelties and contributions of this article can be summarized as follows.

- 1) Rigorous mathematical modeling of the system, where N_P over-the-air (OTA) PIM sources cause nonlinear interference to the RX path, is presented.
- 2) The complete BF method, introduced initially in [17], is formulated and presented in a complete manner, covering both the cancellation processing and the parameter estimation;
- 3) In order to alleviate the modeling and processing complexity, an alternative cancellation scheme leveraging the air-PIM channel coefficients is presented, which, unlike the initial channel coefficient method in [17], can suppress PIM stemming from multiple coexisting air-PIM sources.
- 4) The complexity of each introduced cancellation scheme is studied and quantitatively assessed in terms of floating-point operations (FLOPs).
- 5) Extensive RF measurements are provided with true BS hardware, evidencing that the proposed methods can efficiently cancel air-induced PIM toward the thermal noise floor, and also that the simplified method can essentially match the cancellation performance of the much more complex full BF model in various cases.

The rest of this article is organized as follows. First, in Section II, we provide the modeling fundamentals of the FDD MIMO transceiver under air-induced PIM. Then, in Section III, the proposed digital cancellation methods are described, including the core cancellation processing and the related parameter estimation methods, while also assessing the involved processing complexity. Section IV provides an extensive set of RF measurement-based results. Finally, conclusions are drawn in Section V, while Appendices I and II provide selected technical details related to the BFs and the gradients utilized in the algorithms.

Mathematical Notations: Standard complex baseband signal modeling is adopted throughout this article. Time-dependent variables are denoted with a generic time index t in brackets. The imaginary unit is denoted as j . Boldfaced characters indicate vectors and matrices. The complex conjugate and Hermitian transpose are denoted as $(\cdot)^*$ and $(\cdot)^H$, respectively. Finally, \star denotes the convolution operation.

II. FDD MIMO SYSTEM UNDER AIR-INDUCED PIM

In this section, we introduce the overall system model. We consider a dual-carrier MIMO system with N_t TXs, each

transmitting two CC signals, as shown in Fig. 2. The CC signals are denoted as $x_{1,n}[t]$ and $x_{2,n}[t]$, $n \in \{1, 2, \dots, N_t\}$, such that $x_{1,n}[t]$ are transmitted at frequency f_1 and $x_{2,n}[t]$ are transmitted at f_2 . For analytical purposes and notational convenience, we assume that $f_1 < f_2$. For notational purposes, these frequencies refer to complex baseband equivalent modeling, while the corresponding physical RF frequencies are denoted as $f_{RF,1}$ and $f_{RF,2}$.

Let us assume that there are N_P air-PIM sources in the vicinity of the transceiver such that they are time-aligned at the BB sampling rate of the CCs. The transmit signals propagate through a wireless line-of-sight channel to the PIM sources. Assuming a static frequency flat channel response, the combined signal at the input of the PIM source s can be written as

$$z_s[t] = \sum_{n=1}^{N_t} a_{s,1,n} x_{1,n}[t] e^{j2\pi f_1 t} + a_{s,2,n} x_{2,n}[t] e^{j2\pi f_2 t} \quad (1)$$

$$= \psi_{s,1}[t] e^{j2\pi f_1 t} + \psi_{s,2}[t] e^{j2\pi f_2 t} \quad (2)$$

where $a_{s,1,n}$ and $a_{s,2,n}$ are the channel coefficients corresponding to the channel responses of the CCs 1 and 2, for TX n , toward the s th PIM source.

Following the generalized memory polynomial (GMP) model [39], the s th PIM source output can be written as

$$y_s[t] = \sum_{m=-M_1}^{M_2} \sum_{\substack{p=1 \\ \text{podd}}}^{P_1} g_{s,m,p,0} z_s[t-m] |z_s[t-m]|^{p-1} \\ + \sum_{m=-M_1}^{M_2} \sum_{\substack{p=3 \\ \text{podd}}}^{P_2} \sum_{\substack{k=-K_1 \\ k \neq 0}}^{K_2} g_{s,m,p,k} z_s[t-m] |z_s[t-m-k]|^{p-1} \quad (3)$$

where M_1 and M_2 are the global precursor and postcursor memory depths, respectively, P_1 and P_2 are the polynomial orders considered in the model, K_1 is the envelope lead, K_2 is the envelope lag, and g is the model coefficients.

The PIM source outputs $y_s[t]$ then propagate through wireless channels back to the RX. The received signal $\tilde{r}_{\text{PIM},n}[t]$ at RX n over the whole band can be modeled as

$$\tilde{r}_{\text{PIM},n}[t] = \sum_{s=1}^{N_P} b_{s,n} y_s[t] + \eta[t] \quad (4)$$

where $b_{s,n}$ is the corresponding channel coefficient and $\eta[t]$ is the additive white Gaussian noise. The total received wideband signal can be seen in (5), as shown at the bottom of the next page, where $\tilde{\gamma}$ are the total lumped coefficients of each summed signal ψ , defined in (1) and (2), containing channel coefficients b , the GMP model coefficients g , and the phase shifts stemming from the delays of the memory and envelope lag/lead. Finally, a channel selection filter w_n at the RX limits the bandwidth (BW) of the signal $\tilde{r}_{\text{PIM},n}[t]$ to produce a bandlimited version $r_{\text{PIM},n}[t]$ as

$$r_{\text{PIM},n}[t] = w_n \star \tilde{r}_{\text{PIM},n}[t] + w_n \star \eta[t] \quad (7)$$

where \star denotes the convolution operation. In Section III, building on these fundamental signal models, we concentrate

on the derivation and implementation of the PIM cancellation block in Fig. 2.

III. PROPOSED DIGITAL PIM CANCELLATION METHODS

For generality, the modeling of Section II assumed an arbitrary number of transmit chains N_t and, hence, signals per CC center frequency, and made no assumptions of the corresponding RX frequency. In this section, in order to develop further the closed-form expressions for the received PIM signal, we consider the practical case of $N_t = 2$. Therefore, there are four physical CC signals present in the system, denoted as $x_{1,1}[t]$ and $x_{1,2}[t]$, which occupy frequency f_1 , and $x_{2,1}[t]$ and $x_{2,2}[t]$, which occupy frequency f_2 . In addition, we set the RX frequency to $2f_1 - f_2$, which is the lower third-order intermodulation frequency of the carriers and which, for example, in the case of 5G NR band n3, can fall on the uplink receive band. Accordingly, we will develop models for the received PIM signal for up to polynomial order $P_1 = P_2 = 5$. However, for clarity, it is noted that the modeling results in Section II up to (5) can be used to develop the model for any number of transmitted carriers on any RX frequency. The introduced methods can be used in the PIM canceller block illustrated in Fig. 2. In this section, the individual signals are expressed as time series, where the samples are stacked in vectors. The vector-form TX signals are, therefore, defined as $\mathbf{x}_{1,1}[t] = [x_{1,1}[1], \dots, x_{1,1}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, $\mathbf{x}_{1,2}[t] = [x_{1,2}[1], \dots, x_{1,2}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, which are the signals at frequency f_1 , and $\mathbf{x}_{2,1}[t] = [x_{2,1}[1], \dots, x_{2,1}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, $\mathbf{x}_{2,2}[t] = [x_{2,2}[1], \dots, x_{2,2}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, which are the signals at frequency f_2 . Here, ℓ corresponds to the signal block length.

A. Full BF Method

For brevity, in this section, we neglect the memory effects in our modeling and utilize only the instantaneous model

($m = 0$). This is done without loss of generality, as it is essentially trivial to extend the model to incorporate the past and future samples, since the memory from $m \neq 0$ affects each signal in the same way. With the considered system parameterization, given above, the bandlimited instantaneous received signal at RX n at time instant t , operating on frequency $2f_1 - f_2$, can be written as shown in (6), as shown at the bottom of the page, stemming from (5) and (7), where noise has now been omitted for the notational simplicity.

Using then the definitions in (1) and (2), we can write the instantaneous received PIM signal, as shown in (8), at the bottom of the next page, where the signal is shown up to polynomial order $p = 3$ for brevity. The received PIM signal can then be further defined as expressed in (9), as shown at the bottom of the next page, where the total coefficients of the BFs have been lumped to c and d , such that the coefficients c correspond to BFs, where $k = 0$, and d to BFs where $k \neq 0$. Table IV in Appendix I collects the instantaneous BFs at time instant t for polynomial orders $p = 3$ and $p = 5$ for readers' convenience. It is noted that the number of PIM sources in the system does not affect the number of considered BFs. Conversely, depending on the value of the envelope delay k , the system employs a different number of BFs, depending on whether k is zero or nonzero.

Next, the BFs in Table IV can be stacked into matrix $\boldsymbol{\varphi}[t] \in \mathbb{C}^{\ell \times C}$ each BF occupying a single column. Here, C is the cardinality, i.e., the number of instantaneous BFs in the system. The instantaneous BFs are generated from the input vectors $\mathbf{x}_{1,1}[t]$, $\mathbf{x}_{1,2}[t]$, $\mathbf{x}_{2,1}[t]$, and $\mathbf{x}_{2,2}[t]$ through elementwise multiplications as per Table IV. Incorporating memory into the model can be carried out at his stage by generating the total BF matrix $\Phi[t] = [\boldsymbol{\varphi}[t + M_1], \dots, \boldsymbol{\varphi}[t - M_2]] \in \mathbb{C}^{\ell \times CM}$. Then, the cancellation of the received PIM signal can be written as

$$\mathbf{r}_n[t] = \mathbf{r}_{\text{PIM},n}[t] - \Phi[t]\boldsymbol{\beta} \quad (10)$$

$$\begin{aligned} \tilde{r}_{\text{PIM},n}[t] = & \sum_{s=1}^{N_P} \left(\sum_{m=-M_1}^{M_2} \sum_{\substack{p=1 \\ \text{podd}}}^{P_1} \tilde{\gamma}_{s,m,p,0} \sum_{k_1=0}^{\frac{p-1}{2}} \sum_{k_2=0}^{\frac{p-1}{2}} \binom{p-1}{k_1} \binom{p-1}{k_2} \psi_{s,1}^{k_1+1}[t-m] \psi_{s,1}^{*k_2}[t-m] \psi_{s,2}^{\frac{p-1}{2}-k_1}[t-m] \psi_{s,2}^{*\frac{p-1}{2}-k_2}[t-m] \right. \\ & \times e^{j2\pi((k_1-k_2+1)f_1+(k_2-k_1)f_2)t} + \psi_{s,1}^{k_1}[t-m] \psi_{s,1}^{*k_2}[t-m] \psi_{s,2}^{\frac{p+1}{2}-k_1}[t-m] \psi_{s,2}^{*\frac{p-1}{2}-k_2}[t-m] e^{j2\pi((k_1-k_2)f_1+(k_2-k_1+1)f_2)t} \\ & + \sum_{m=-M_1}^{M_2} \sum_{\substack{p=3 \\ \text{podd}}}^{P_2} \sum_{\substack{k=-K_1 \\ k \neq 0}}^{K_2} \tilde{\gamma}_{s,m,p,k} \sum_{k_3=0}^{\frac{p-1}{2}} \sum_{k_4=0}^{\frac{p-1}{2}} \binom{p-1}{k_3} \binom{p-1}{k_4} \psi_{s,1}^{k_3}[t-m] \psi_{s,1}^{*k_4}[t-m] \psi_{s,2}^{\frac{p-1}{2}-k_3}[t-m] \psi_{s,2}^{*\frac{p-1}{2}-k_4}[t-m] \\ & \times e^{j2\pi((k_1-k_2+1)f_1+(k_2-k_1)f_2)t} \\ & \left. + \psi_{s,2}[t-m] \psi_{s,1}^{k_3}[t-m] \psi_{s,1}^{*k_4}[t-m] \psi_{s,2}^{\frac{p-1}{2}-k_3}[t-m] \psi_{s,2}^{*\frac{p-1}{2}-k_4}[t-m] e^{j2\pi((k_1-k_2)f_1+(k_2-k_1+1)f_2)t} \right) + \eta[t] \quad (5) \end{aligned}$$

$$\begin{aligned} r_{\text{PIM},n}[t] = & \sum_{s=1}^{N_P} \left(\gamma_{s,0,3,0} \psi_{s,1}^2[t] \psi_{s,2}^*[t] + \gamma_{s,0,5,0} (\psi_{s,1}^2[t] \psi_{s,2}^*[t] |\psi_{s,1}[t]|^2 + \psi_{s,1}^2[t] \psi_{s,2}^*[t] |\psi_{s,2}[t]|^2) \right. \\ & + \dots + \sum_{\substack{k=-K_1 \\ k \neq 0}}^{K_2} \gamma_{s,0,3,k} \psi_{s,1}[t] \psi_{s,1}[t-k] \psi_{s,2}^*[t-k] + \gamma_{s,0,5,k} (\psi_{s,1}[t] \psi_{s,1}[t-k] \psi_{s,2}^*[t-k] |\psi_{s,1}[t-k]|^2 \\ & \left. + \psi_{s,1}[t] \psi_{s,1}[t-k] \psi_{s,2}^*[t-k] |\psi_{s,2}[t-k]|^2 + \psi_{s,2}[t] \psi_{s,1}^2[t-k] \psi_{s,2}^*[t-k]) + \dots \right) \quad (6) \end{aligned}$$

where $\mathbf{r}_n[t] = [r_n[1], \dots, r_n[\ell]]^T \in \mathbb{C}^{\ell \times 1}$ is the n th RX signal with the PIM suppressed, $\mathbf{r}_{\text{PIM},n}[t] = [r_{\text{PIM},n}[1], \dots, r_{\text{PIM},n}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$ is the n th RX signal stacked in to a vector, and $\boldsymbol{\beta} \in \mathbb{C}^{CM \times 1}$ contains the estimates of the coefficients c and d of each BF shown in (9). For future reference, following (8) and (9), the coefficients c —corresponding to the instantaneous third-order ($p = 3$) BFs of the model—can be explicitly expressed as

$$\begin{aligned} c_1 &= \sum_{s=1}^{N_p} \gamma_{s,0,3,0} a_{s,1,1}^2 a_{s,2,1}^* \\ c_2 &= \sum_{s=1}^{N_p} 2\gamma_{s,0,3,0} a_{s,1,1} a_{s,1,2} a_{s,2,1}^* \\ c_3 &= \sum_{s=1}^{N_p} \gamma_{s,0,3,0} a_{s,1,2}^2 a_{s,2,1}^* \\ c_4 &= \sum_{s=1}^{N_p} \gamma_{s,0,3,0} a_{s,1,1}^2 a_{s,2,2}^* \\ c_5 &= \sum_{s=1}^{N_p} 2\gamma_{s,0,3,0} a_{s,1,1} a_{s,1,2} a_{s,2,2}^* \\ c_6 &= \sum_{s=1}^{N_p} \gamma_{s,0,3,0} a_{s,1,2}^2 a_{s,2,2}^*. \end{aligned} \quad (11)$$

A block diagram of the full BF scheme is illustrated in Fig. 3. Furthermore, due to the linear-in-parameters nature of the

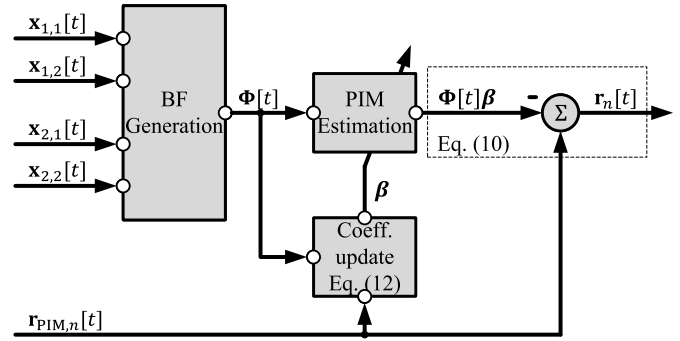


Fig. 3. Block diagram of the cancellation scheme of the full BF model, as described in Section III-A.

model utilizing the BFs of Table IV, the coefficients $\boldsymbol{\beta}$ can be estimated using least squares (LS) fitting as

$$\boldsymbol{\beta} = (\Phi^H[t]\Phi[t])^{-1} \Phi^H[t] \mathbf{r}_{\text{PIM},n}[t]. \quad (12)$$

The required steps to cancel PIM with the full BF method are shown and summarized in Algorithm 1 for a given block of received signal samples. In a static environment, the recalculation of the coefficient vector $\boldsymbol{\beta}$ is not always necessary for each processing block.

Since the complexity of the cancellation engine in (10) and especially that of the coefficients estimation in (12) are dependent on the number of BFs in the system, the complexity of the above proposed scheme is extremely high, especially when memory is considered in the model. Inspired by this, in

$$\begin{aligned} r_{\text{PIM},n}[t] &= \sum_{s=1}^{N_p} \left(\gamma_{s,0,3,0} (a_{s,1,1} a_{s,2,1}^* x_{1,1}^2[t] x_{2,1}^*[t] + 2a_{s,1,1} a_{s,1,2} a_{s,2,1}^* x_{1,1}[t] x_{1,2}[t] x_{2,1}^*[t] + a_{s,1,2} a_{s,2,1}^* x_{1,2}^2[t] x_{2,1}^*[t] \right. \\ &\quad + a_{s,1,1} a_{s,2,2}^* x_{1,1}^2[t] x_{2,2}^*[t] + 2a_{s,1,1} a_{s,1,2} a_{s,2,2}^* x_{1,1}[t] x_{1,2}[t] x_{2,2}^*[t] + a_{s,1,2} a_{s,2,2}^* x_{1,2}^2[t] x_{2,2}^*[t]) + \dots \\ &\quad + \dots + \sum_{\substack{k=-K_1 \\ k \neq 0}}^{K_2} \gamma_{s,0,3,k} (a_{s,1,1}^2 a_{s,2,1}^* x_{1,1}[t] x_{1,1}[t-k] x_{2,1}^*[t-k] \\ &\quad + a_{s,1,1}^2 a_{s,2,2}^* x_{1,1}[t] x_{1,1}[t-k] x_{2,2}^*[t-k] + a_{s,1,2}^2 a_{s,2,1}^* x_{1,2}[t] x_{1,2}[t-k] x_{2,1}^*[t-k] \\ &\quad + a_{s,1,1} a_{s,1,2} a_{s,2,1}^* x_{1,1}[t] x_{1,2}[t-k] x_{2,1}^*[t-k] \\ &\quad + a_{s,1,1} a_{s,1,2} a_{s,2,2}^* x_{1,1}[t] x_{1,2}[t-k] x_{2,2}^*[t-k] + a_{s,1,2}^2 a_{s,2,2}^* x_{1,2}[t] x_{1,2}[t-k] x_{2,2}^*[t-k] \\ &\quad \left. + a_{s,1,1} a_{s,1,2} a_{s,2,1}^* x_{1,2}[t] x_{1,1}[t-k] x_{2,1}^*[t-k] + a_{s,1,1} a_{s,1,2} a_{s,2,2}^* x_{1,2}[t] x_{1,1}[t-k] x_{2,2}^*[t-k] + \dots \right) \end{aligned} \quad (8)$$

$$\begin{aligned} r_{\text{PIM},n}[t] &= c_1 x_{1,1}^2[t] x_{2,1}^*[t] + c_2 x_{1,1}[t] x_{1,2}[t] x_{2,1}^*[t] \\ &\quad + c_3 x_{1,2}^2[t] x_{2,1}^*[t] + c_4 x_{1,1}^2[t] x_{2,2}^*[t] + c_5 x_{1,1}[t] x_{1,2}[t] x_{2,2}^*[t] + c_6 x_{1,2}^2[t] x_{2,2}^*[t] \\ &\quad + \dots + \sum_{\substack{k=-K_1 \\ k \neq 0}}^{K_2} d_1 x_{1,1}[t] x_{1,1}[t-k] x_{2,1}^*[t-k] + d_2 x_{1,1}[t] x_{1,1}[t-k] x_{2,2}^*[t-k] \\ &\quad + d_3 x_{1,2}[t] x_{1,2}[t-k] x_{2,1}^*[t-k] + d_4 x_{1,1}[t] x_{1,2}[t-k] x_{2,1}^*[t-k] \\ &\quad + d_5 x_{1,1}[t] x_{1,2}[t-k] x_{2,2}^*[t-k] + d_6 x_{1,2}[t] x_{1,2}[t-k] x_{2,2}^*[t-k] + d_7 x_{1,2}[t] x_{1,1}[t-k] x_{2,1}^*[t-k] \\ &\quad + d_8 x_{1,2}[t] x_{1,1}[t-k] x_{2,2}^*[t-k] + \dots \end{aligned} \quad (9)$$

Algorithm 1 Full BF Method for PIM Estimation and Cancellation

```

1: Inputs:  $\mathbf{x}_{1,1}[t]$ ,  $\mathbf{x}_{1,2}[t]$ ,  $\mathbf{x}_{2,1}[t]$ ,  $\mathbf{x}_{2,2}[t]$ ,  $\mathbf{r}_{\text{PIM},n}[t]$ 
2: while cancelling do
3:   Calculate  $\Phi[t]$  as per Table IV in Appendix I
4:   Calculate  $\boldsymbol{\beta}$  as per Eq. (12)
5:   Calculate  $\Phi[t]\boldsymbol{\beta}$ 
6:   Calculate  $\mathbf{r}_n[t]$  as per Eq. (10)
7: end while

```

Sections III-B and III-C, we propose alternative cancellation schemes with much reduced complexity, leveraging on the knowledge of the channel coefficients $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, and $a_{s,2,2}$.

B. Channel Coefficient Method: Single PIM Source

Next, we consider the special case of having only a single air-PIM source in the vicinity of the transceiver, i.e., $N_P = 1$. The cancellation scheme covered in this subsection has been initially introduced in our early work in [17], while it is also stated here for the sake of completeness and readability. To this end, the computational complexity of the full BF method introduced in Section III-A can be alleviated by summing the signals on the same frequencies, as indicated already by (1), where the relative phases and amplitudes of the superimposed signals are taken into account. Consequently, we define two aggregated signals $\mathbf{v}_1[t] = [v_1[1], \dots, v_1[\ell]]^T \in \mathbb{C}^{\ell \times 1}$ and $\mathbf{v}_2[t] = [v_2[1], \dots, v_2[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, calculated as

$$\mathbf{v}_1[t] = q_1 \mathbf{x}_{1,1}[t] + \mathbf{x}_{1,2}[t] \quad (13)$$

$$\mathbf{v}_2[t] = q_2 \mathbf{x}_{2,1}[t] + \mathbf{x}_{2,2}[t] \quad (14)$$

where the complex multipliers are defined as $q_1 = a_{1,1,1}/a_{1,1,2}$ and $q_2 = a_{1,2,1}/a_{1,2,2}$, which account for the relative phase and magnitude of the signals on the same frequency. From (8), (9), and (11), and when $N_P = 1$, we can see that the coefficients q can then be obtained from the coefficients c_i , $i = 1, 2, \dots, 6$, as follows:

$$q_1 = \frac{2c_1}{c_2} = \frac{c_2}{2c_3} = \pm \sqrt{\frac{c_1}{c_3}} = \frac{2c_4}{c_5} = \frac{c_5}{2c_6} = \pm \sqrt{\frac{c_4}{c_6}} \quad (15)$$

$$q_2 = \frac{c_1^*}{c_4^*} = \frac{c_2^*}{c_5^*} = \frac{c_3^*}{c_6^*}. \quad (16)$$

This indicates that we only require a third-order instantaneous model of the received signal to determine the q coefficients. For improved accuracy in the implementation of the algorithm, the coefficients q_1 and q_2 can be defined as arithmetic means of the shown expressions in (15) and (16). We can then construct the BF matrix $\Phi[t]$ of (10) using the BFs expressed in Table I by doing elementwise multiplication using vectors $\mathbf{v}_1[t]$ and $\mathbf{v}_2[t]$. The amount of the incorporated BFs is now significantly reduced compared to the full BF method, indicating largely reduced computational complexity.

The coefficients c in the model of the received signal required by the computation of the q coefficients can be found, e.g., via LS fitting. To this end, let us define vector $\mathbf{c} = [c_1, \dots, c_6]^T \in \mathbb{C}^{6 \times 1}$ containing the instantaneous third-order

Algorithm 2 Channel Coefficient Method for PIM Estimation and Cancellation, and Single PIM Source

```

1: Inputs:  $\mathbf{x}_{1,1}[t]$ ,  $\mathbf{x}_{1,2}[t]$ ,  $\mathbf{x}_{2,1}[t]$ ,  $\mathbf{x}_{2,2}[t]$ ,  $\mathbf{r}_{\text{PIM},n}[t]$ 
2: while cancelling do
3:   Calculate  $\mathbf{c}$  as per Eq. (17)
4:   Calculate  $q_1$  and  $q_2$  as per Eqs. (15), (16)
5:   Calculate  $\mathbf{v}_1[t]$  and  $\mathbf{v}_2[t]$  as per Eqs. (13), (14)
6:   Calculate  $\Phi[t]$  as per Table I
7:   Calculate  $\boldsymbol{\beta}$  as per Eq. (12)
8:   Calculate  $\Phi[t]\boldsymbol{\beta}$ 
9:   Calculate  $\mathbf{r}_n[t]$  as per Eq. (10)
10: end while

```

TABLE I

INSTANTANEOUS ($m = 0$) BFs OF THE CHANNEL COEFFICIENT METHOD UP TO POLYNOMIAL ORDER $p = 5$

p	Unique Inst. BFs	Number of BFs	
		$k = 0$	$k \neq 0$
3	$v_1[t]v_1[t-k]v_2^*[t-k]$	1	1
5	$v_1[t]v_1[t-k]v_2^*[t-k] v_1[t-k] ^2$	2	3
	$v_2[t]v_1^*[t-k]v_2^*[t-k]$		

BF coefficients of (11). Then, we can find the LS estimate of the required c coefficients as

$$\mathbf{c} = (\boldsymbol{\theta}^H[t]\boldsymbol{\theta}[t])^{-1}\boldsymbol{\theta}^H[t]\mathbf{r}_{\text{PIM},n}[t] \quad (17)$$

where $\boldsymbol{\theta}[t] \in \mathbb{C}^{\ell \times 6}$ contains the instantaneous, zero envelope lag ($k = 0$) third-order BFs of Table IV, stacked to a matrix, such that each BF occupies a single column. This means that the channel coefficient method requires the coefficient estimation to be carried out twice, rather than once as is the case with the full BF modeling. Despite this, the complexity reduction from this technique in the main path and learning processing is substantial. This will be quantified explicitly in Section III-D. Algorithm 2 compiles and summarizes the steps required by the single PIM source channel coefficient method. As in the full BF method, the recalculation of the LS-based variables $\boldsymbol{\beta}$ and \mathbf{c} (and, thus, also q_1 and q_2) is not necessary on every iteration in a static environment.

C. Channel Coefficient Method: Multiple PIM Sources

Even though the method introduced in Section III-B offers a great reduction in computational complexity compared to the full BF method, its usability is restricted to cases with a single dominant air-PIM source. To this end, in this section, we introduce a more general cancellation scheme with reduced main path complexity compared to the full BF model, which can handle any number of PIM sources in the vicinity of the transceiver. The basic concept of this cancellation scheme is illustrated as a block diagram in Fig. 4.

From the definition of the coefficients c in (11), it stems that the channel coefficient method of Section III-B can only handle a single PIM source—as already noted—since the above coefficients describing the situation with multiple PIM sources contain summations. On the other hand, we can now

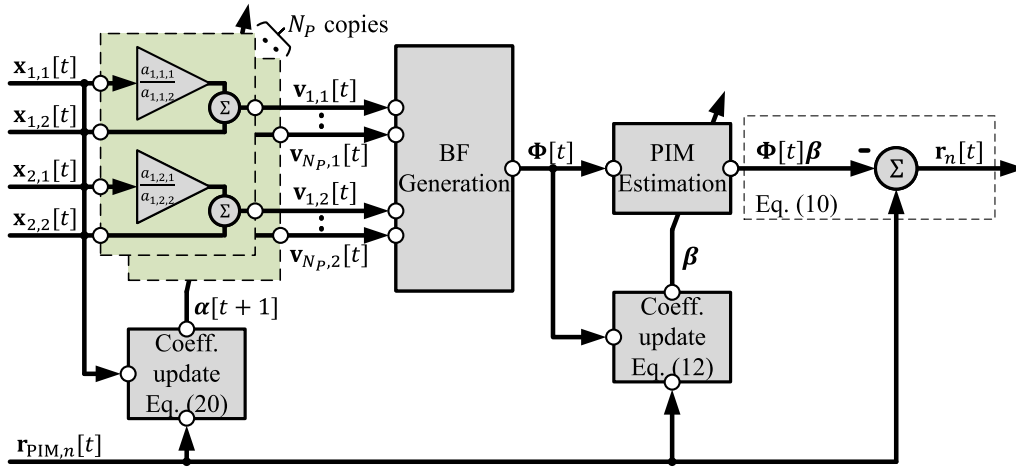


Fig. 4. Block diagram of the channel coefficient model-based cancellation scheme, as described in Section III-C.

build an estimate of the received PIM signal by utilizing an instantaneous third-order model, expressed as

$$\begin{aligned} \hat{r}[t] = & c_1 x_{1,1}^2[t] x_{2,1}^* [t] + c_2 x_{1,1}[t] x_{1,2}[t] x_{2,1}^* [t] \\ & + c_3 x_{1,2}^2[t] x_{2,1}^* [t] + c_4 x_{1,1}^2[t] x_{2,2}^* [t] \\ & + c_5 x_{1,1}[t] x_{1,2}[t] x_{2,2}^* [t] + c_6 x_{1,2}^2[t] x_{2,2}^* [t] \end{aligned} \quad (18)$$

where the estimated PIM signal $\hat{r}[t]$ is dependent on all the channel coefficients $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, and $a_{s,2,2}$, and the lumped coefficients $\gamma_{s,0,3,0}$. The instantaneous third-order estimates at each time instant t can be stacked to a vector $\hat{\mathbf{r}} = [\hat{r}[1], \dots, \hat{r}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$. Now, we can define an error signal $\mathbf{e}[t] \in \mathbb{C}^{\ell \times 1}$ as

$$\mathbf{e}[t] = \mathbf{r}_{\text{PIM},n}[t] - \hat{\mathbf{r}}[t] \quad (19)$$

which measures how well our model matches the measured data. We can further define a loss function for the system as $L[t] = \|\mathbf{e}[t]\|_2^2$, which we aim to minimize by finding optimal values for the channel coefficients $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, and $a_{s,2,2}$, and for the lumped coefficients $\gamma_{s,0,3,0}$. For solving this, we employ the Gauss–Newton iterative method, as the problem is now *nonlinear-in-parameters* in its nature. To this end, by stacking the unknown parameters to vector $\mathbf{a}[t] \in \mathbb{C}^{5N_p \times 1}$, the parameter update following the damped Gauss–Newton method, also known as the Levenberg–Marquardt algorithm, can be written as [40]

$$\mathbf{a}[t+1] = \mathbf{a}[t] + (\mathbf{J}^H[t] \mathbf{J}[t] + \lambda[t] \mathbf{I})^{-1} \mathbf{J}^H[t] \mathbf{e}[t] \quad (20)$$

where $\mathbf{J}[t] \in \mathbb{C}^{\ell \times 5N_p}$ is the Jacobian matrix of the function $\hat{r}[t]$, $\lambda[t]$ is the diagonal loading factor, and \mathbf{I} is the appropriately sized identity matrix.

The Jacobian matrix $\mathbf{J}[t]$ is defined by the partial derivatives of the function $\hat{r}[t]$ w.r.t. all the parameters, namely, $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, $a_{s,2,2}$, and $\gamma_{s,0,1,0}$, for $s = 1, 2, \dots, N_p$. Due to the definition of the instantaneous estimate of the signal $\hat{r}[t]$, we can define the total Jacobian $\mathbf{J}[t]$ using submatrices $\mathbf{J}_s[t] \in \mathbb{C}^{\ell \times 5}$, as

$$\mathbf{J}[t] = [\mathbf{J}_1[t] \mathbf{J}_2[t], \dots, \mathbf{J}_{N_p}[t]]. \quad (21)$$

Now, each submatrix $\mathbf{J}_s[t]$ can be written as

$$\mathbf{J}_s[t] = \begin{bmatrix} \frac{\partial \hat{r}[1]}{\partial a_{s,1,1}} & \frac{\partial \hat{r}[1]}{\partial a_{s,1,2}} & \frac{\partial \hat{r}[1]}{\partial a_{s,2,1}^*} & \frac{\partial \hat{r}[1]}{\partial a_{s,2,2}^*} & \frac{\partial \hat{r}[1]}{\partial \gamma_{s,0,1,0}} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \hat{r}[\ell]}{\partial a_{s,1,1}} & \frac{\partial \hat{r}[\ell]}{\partial a_{s,1,2}} & \frac{\partial \hat{r}[\ell]}{\partial a_{s,2,1}^*} & \frac{\partial \hat{r}[\ell]}{\partial a_{s,2,2}^*} & \frac{\partial \hat{r}[\ell]}{\partial \gamma_{s,0,1,0}} \end{bmatrix} \quad (22)$$

where the partial derivatives are defined and expressed, as shown in Appendix II.

Using all the channel coefficients in $\mathbf{a}[t]$, we define N_p signal pairs of the form $\mathbf{v}_{s,1}[t] = [v_{s,1}[1], \dots, v_{s,1}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$ and $\mathbf{v}_{s,2}[t] = [v_{s,2}[1], \dots, v_{s,2}[\ell]]^T \in \mathbb{C}^{\ell \times 1}$, expressed as

$$\mathbf{v}_{s,1}[t] = (a_{s,1,1}/a_{s,1,2}) \mathbf{x}_{1,1}[t] + \mathbf{x}_{1,2}[t] \quad (23)$$

$$\mathbf{v}_{s,2}[t] = (a_{s,2,1}/a_{s,2,2}) \mathbf{x}_{2,1}[t] + \mathbf{x}_{2,2}[t] \quad (24)$$

similarly as in the single PIM source case in Section III-B. The BF matrix $\Phi[t]$ is then generated using the signals $\mathbf{v}_{s,1}[t]$ and $\mathbf{v}_{s,2}[t]$, as in Table I, where $v_1[t]$ corresponds to each $v_{s,1}[t]$, and $v_2[t]$ to each $v_{s,2}[t]$ for $s = 1, 2, \dots, N_p$. Then, the cancellation of the PIM is carried out as described in (10) and (12). The steps of the channel coefficient method with multiple PIM sources are shown in Algorithm 3. Again, the recalculation of vector β on every iteration is not necessary for a static environment. Now, the BF matrix $\Phi[t]$ consists of N_p independent copies of the BFs shown and summarized in Table I, i.e., there are N_p third-order BFs, and $2N_p$ or $3N_p$ fifth-order BFs, depending on whether k is zero or nonzero. Therefore, the main path complexity of the proposed channel coefficient method with multiple PIM sources has slightly increased compared to the single PIM source method, but we still have a considerable complexity reduction compared to the full BF method when N_p is moderate. This is addressed and quantified more concretely in Section III-D.

D. Computational Complexity Analysis

Here, we investigate the computational complexity of the aforementioned cancellation schemes in terms of FLOPs.

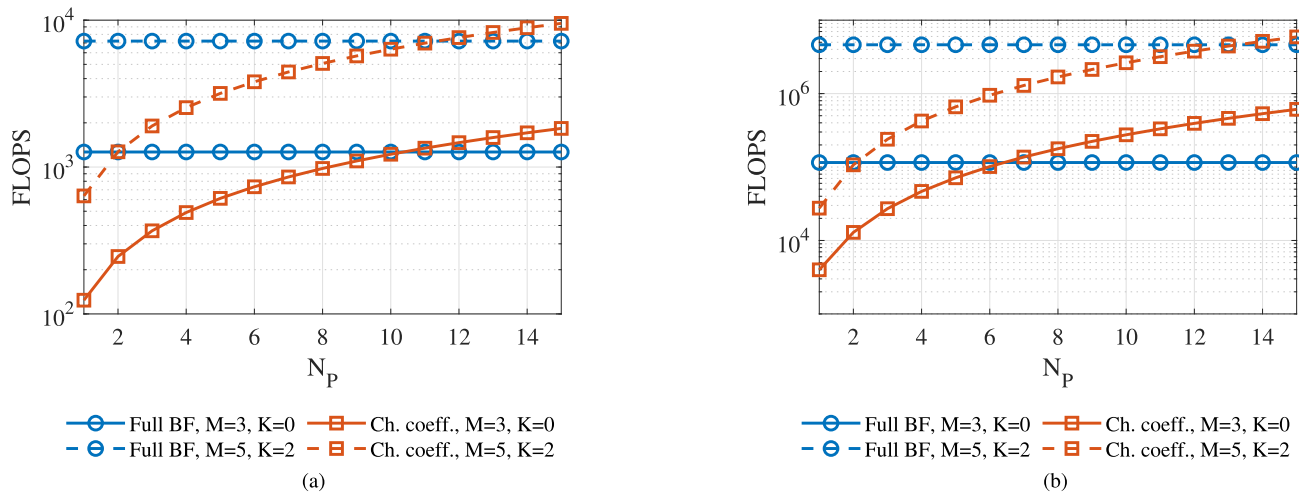


Fig. 5. Computational complexities in terms of FLOPs of the full BF and channel coefficient estimation methods for the main path processing and parameter updates, assuming fifth-order cancellers. The complexity is calculated per processed output sample in the main path and per block of data in the learning. The learning in the channel coefficient model includes the complexity of the channel coefficient estimation. The iterations for Gauss–Newton to converge, i_c , is set to 10. (a) Main path. (b) Learning.

TABLE II
COMPUTATIONAL COMPLEXITIES OF THE INVESTIGATED CANCELLATION SCHEMES AS FLOPS PER SAMPLE IN THE MAIN PATH AND PER BLOCK IN THE LEARNING

Operation	Full BF (Sec. III-A)	Ch. Coeff. (Sec. III-B)	Ch. Coeff. (Sec. III-C)
Main Path	$448KM + 320M + 408K + 306$	$(32KM + 24M + 72K + 50)N_P + 2$	
Learning			
β	$(25088K^2 + 35840K + 12800)M^2$	$(128K^2 + 192K + 72)M^2N_P^2$	
Ch. coeffs.	N/A	342	$(200N_P^2 + 96N_P + 38)i_c$

Algorithm 3 Channel Coefficient Method for PIM Estimation and Cancellation, and Multiple PIM Sources

- 1: **Require:** Initialize $\alpha[t]$, $\lambda[t]$
- 2: **Inputs:** $\mathbf{x}_{1,1}[t]$, $\mathbf{x}_{1,2}[t]$, $\mathbf{x}_{2,1}[t]$, $\mathbf{x}_{2,2}[t]$, $\mathbf{r}_{\text{PIM},n}[t]$
- 3: **while** cancelling **do**
- 4: Initialize $\Phi[t]$
- 5: **for** $s = 1$ to N_P **do**
- 6: Obtain $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, and $a_{s,2,2}$ from $\alpha[t]$
- 7: Calculate $\mathbf{v}_{s,1}[t]$ and $\mathbf{v}_{s,2}[t]$ as per Eqs. (23), (24)
- 8: Append BFs to $\Phi[t]$ as per Table I
- 9: **end for**
- 10: Calculate β as per Eq. (12)
- 11: Calculate $\Phi[t]\beta$
- 12: Calculate $\mathbf{r}_n[t]$ as per Eq. (10)
- 13: Calculate $\mathbf{e}[t]$ as per Eq. (19)
- 14: Calculate $\mathbf{J}[t]$ as per Eqs. (21), (22)
- 15: Update $\alpha[t]$ as per Eq. (20)
- 16: Update $\lambda[t]$
- 17: **end while**

The computational complexities are shown in Table II, where $M = M_1 + M_2 + 1$ denotes the total memory, $K = K_1 + K_2$ refers to the total envelope lead/lag, and i_c denotes the number of iterations it takes the Gauss–Newton algorithm in (20) to converge to a solution. The results are shown *per output sample* in the main path and *per processed data block* in the learning such that the complexity is normalized with the block size. To obtain the presented results, it is assumed that

a complex multiplication takes eight FLOPs, and a complex addition takes two FLOPs. In addition, it is assumed that only the generation of the instantaneous BFs is necessary per iteration. Furthermore, QR decomposition is assumed to be used in the LS and Gauss–Newton algorithms. The complexity of the QR decomposition can be estimated to be $8n^2$ FLOPs per block, where n is the number of columns in the matrices $\Phi[t]$ or $\mathbf{J}[t]$ [41]. For improved accuracy, the q coefficients in (15) and (16) are calculated as the means of the shown definitions, excluding the terms with square roots, as these are resource-hungry operations.

Fig. 5 shows the main path and learning complexities of the full BF method and channel coefficient method of Section III-C, with two different system parameterizations and with various numbers of PIM sources N_P . Here, we have set the number of iterations for the Gauss–Newton algorithm to converge as $i_c = 10$. It can be seen from Fig. 5 and Table II that both the main path and learning complexities of the full BF method are invariant w.r.t. the number of PIM sources N_P . On the other hand, the complexity of both the main path and learning of the channel coefficient method increases with N_P . Since the complexity of the generation of the instantaneous BFs increases with N_P , the main path complexity of the channel coefficient method surpasses that of the full BF model when sufficiently many PIM sources are considered; however, this occurs only when $N_P > 10$ for both considered parameterizations. Likewise, the complexity of the learning in the channel coefficient methods increases with N_P . With the lower parameterization of $M = 3$ and $K = 0$, the full BF model complexity is surpassed when $N_P > 6$, and with

$M = 5$ and $K = 2$ when $N_P > 13$. These results indicate the superiority of the proposed alternative methods in terms of computational complexity when a moderate number of significant PIM sources is considered in the system. It is noted here that, if many significant sources of PIM are located near the transceiver, relocating the transceiver or the corresponding antenna system is likely the optimal solution.

Moreover, by setting $N_P = 1$ in Table II, we can see that the complexity to determine the channel coefficients in the multiple PIM source method from Section III-C is $334i_c$ FLOPs. Therefore, the proposed method of Section III-C is approximately as complex as the single PIM source method from Section III-B with a single PIM source per iteration, while also being a more general solution, capable of handling any number of PIM sources. However, since the multiple PIM source method requires multiple iterations to reach the optimal solution, there is a tradeoff between speed and complexity when only a single dominant PIM source is present in the system.

It is worth noting that changes in signal BW do not affect the relative computational complexities presented in this section of any of the models, as long as the parameterization of the algorithms is kept unaltered, since the provided values are normalized per sample and per processed block. Yet, if a high signal BW would require higher memory depth (M) in the modeling of the nonlinear response of the canceller than what is shown in Fig. 5, the relative complexity savings from the introduced channel coefficient methods would be even more substantial.

IV. RF MEASUREMENTS AND RESULTS

In order to validate the introduced PIM cancellation methods, a set of RF measurements was conducted using real-life cellular BS hardware. We will first describe the setup used for the measurements and then present the PIM cancellation results with one, two, and three air-PIM sources while considering also two transmit signal BWs for the CCs.

A. Measurement Setup

The RF measurement setup is shown in Fig. 6(a). All the measurements were conducted within an anechoic chamber, where a dual-TX/RX BS transmits and receives in FDD operation, transmitting four physical carrier waveforms, as described in Section III. The TX RF center frequencies are set to 5G NR band n3 frequencies of $f_{RF,1} = 1819.0$ MHz and $f_{RF,2} = 1866.5$ MHz, while the RX operates at the n3 receive band, at 1771.5 MHz, which is the lower third-order intermodulation frequency ($2f_{RF,1} - f_{RF,2}$) of the TX frequencies. A PC outside the anechoic chamber is used to control the BS and collect data for postprocessing. In the measurements, 5G NR standard-compliant cyclic prefix orthogonal frequency division multiplexing (CP-OFDM) waveforms of the chosen carrier BWs of 5 and 20 MHz are utilized. The baseband sampling rates of the carriers and the RX signal are 7.68 and 30.72 MHz, respectively, for the two considered BWs. In the postprocessing, where the PIM cancellation takes place, the TX and RX signals are oversampled by a factor of 16 for the BF generation and synchronization tasks. The parameter

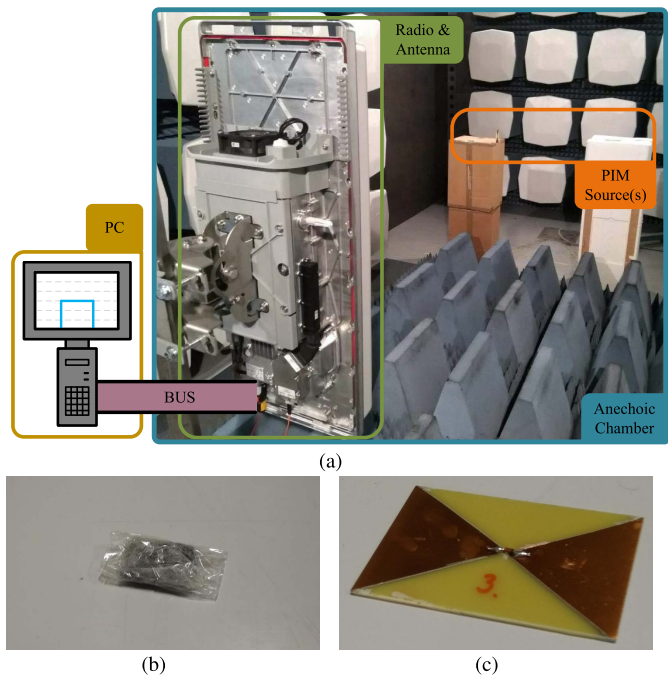


Fig. 6. (a) RF measurement setup, where the important parts of the setup are highlighted. The PIM sources used in the experiments are also shown: (b) piece of off-the-shelf steel wool and (c) purpose-made ribbon antenna with a diode in the middle.

estimation and the actual cancellation are carried out at the original baseband sampling rates. In the measurements, the obtained RX signal data vectors, captured over 10 ms, are always divided into two: the parameter estimation is carried out using the first half, and the cancellation is performed and evaluated in the second half.

The directive antennas of the BS with an antenna gain of around 16 dBi were aimed toward the tested PIM source(s) such that the TX signals hit the PIM source(s), undergo a nonlinear transformation, and are reflected toward the BS RX paths. Due to the high transmit power required to generate meaningful or observable levels of air-induced PIM, the TX signals are predistorted with a built-in DPD solution to avoid nonlinear distortion stemming from the PAs. The TX signals then fulfill the spectral emission limits set by the 5G NR standards [2]. It is also noted here that, since the intermodulation distortion considered in this work is generated outside the TX, there are no limits for such emissions in the current 5G NR standards. Specifically, there are additional emission requirements and limits defined in [2] for BS TXs in order to protect the BS's own and/or other colocated RXs. These requirements are, however, defined either for conducted measurements or for OTA measurements in a pure chamber and, hence, do not consider the air-PIM induced by the built environment around the BS antenna system in true deployment sites, which is the air-induced PIM problem studied in this article.

In the measurement arrangement, the PIM signals appear essentially synchronously from the PIM sources, as they are within the chamber. As a concrete example, considering a 5-MHz signal sampled at $f_s = 7.68$ MHz, the distance

TABLE III
ADDITIONAL INSTANTANEOUS BFs FOR THE NINTH-ORDER CHANNEL
COEFFICIENT MODEL

p	Inst. BFs
7	$v_1^2[t]v_2^*[t] v_1[t] ^4 v_1^2 v_2^*[t] v_2[t] ^4$ $v_1^2[t]v_2^*[t] v_1[t] ^2 v_2[t] ^2$
9	$v_1^2[t]v_2^*[t] v_1[t] ^6 v_1^2 v_2^*[t] v_2[t] ^6$ $v_1^2[t]v_2^*[t] v_1[t] ^4 v_2[t] ^2 v_1^2[t]v_2^*[t] v_1[t] ^2 v_2[t] ^4$

difference between the PIM sources can be approximately 39 m, measured along the line-of-sight of the transceiver, for them to appear within the same sample, which is far greater than the dimensions of the anechoic chamber where the PIM sources lie. This distance is additionally so large that any potential sources close to the 39 m limit would probably not induce any significant PIM distortion since, in general, the larger the distance to the potential PIM source, the smaller the observed PIM on the RX band will be. Two different types of PIM sources were utilized in the measurements: standard off-the-shelf steel wool and purpose-made bow tie antennas, where the nonlinear response is achieved with a diode soldered to the middle of the antenna. The PIM sources are shown in Fig. 6(b) and (c). In the following, results of only RX path 1 are shown, for presentation simplicity, but the results from RX path 2 are similar.

B. Single PIM Source

First, let us examine the cancellation capabilities of the proposed PIM cancellation schemes with a single air-PIM source, i.e., $N_P = 1$, where the PIM source in question is placed around 2.5 m from the BS unit. In Fig. 7, the achieved PIM cancellation is plotted with various TX carrier powers utilizing 5-MHz carriers with the parameterization of $M_1 = M_2 = 2$ and $K_1 = K_2 = 1$. The cancellers adopt the fifth-order modeling introduced in Section III; therefore, the full BF model employs 760 BFs in the matrix $\Phi[t]$, while the channel coefficient methods employ only 65 BFs. In addition, the performance of a ninth-order canceller employing the channel coefficient method of Section III-C is shown. Furthermore, Fig. 8 shows the power spectral densities (PSDs) of the involved signals in the +31 dBm TX power case. It can be seen that all the fifth-order cancellers achieve the same level of cancellation with all the TX powers, indicating that the drastic reduction in complexity from the channel coefficient models does not compromise the cancellation capability. All of the fifth-order models mitigate the PIM all the way to the noise floor with TX powers up to +27 dBm. The residual PIM with fifth-order models then starts increasing with increased TX power, reaching a level of up to 2.3 dB at +31-dBm TX power, with the cancellation being 17.3 dB and the maximum being 19.6 dB in this case. These results are well in line with the results from our previous work in [17], where cancellation of up to 20 dB was evidenced in a similar setup.

The near 3 dB difference in cancellation compared to the one reported in our previous work [17] can be explained by the sensitivity of the setup, particularly the exact orientation and

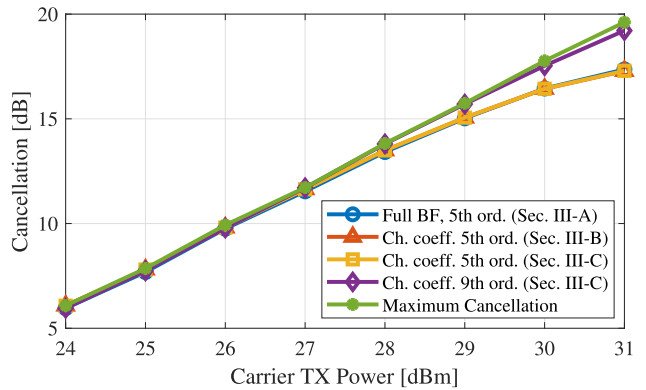


Fig. 7. Power sweep with 5-MHz carriers, steel wool as PIM source, $M_1 = M_2 = 2$, $K_1 = K_2 = 1$.

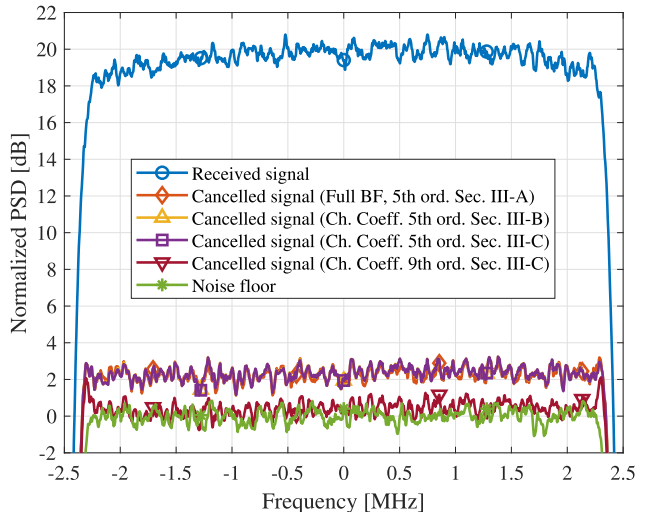


Fig. 8. Cancellation of the 5-MHz carrier PIM from steel wool with +31-dBm TX power, $M_1 = M_2 = 2$, and $K_1 = K_2 = 1$.

location of the steel wool PIM source w.r.t. the BS. Therefore, in addition to the fifth-order modeling, Figs. 7 and 8 demonstrate the higher cancellation fidelity obtained by the ninth-order model with the channel coefficient methods. The ninth-order model contains additional memory polynomial BFs ($k = 0$) up to polynomial order $P_1 = 9$, with the additional instantaneous BFs being shown in Table III. With the added BFs, the cancellation at +31 dBm TX power is now 19.3 dB, merely 0.6 dB from the noise floor. This added performance comes with the price of only adding seven new instantaneous BFs, which with the considered five taps of memory adds up to 35 new BFs, totaling then 100 BFs in the matrix $\Phi[t]$. In order to reach this level of accuracy in the model with the full BF modeling, BFs in the order of a thousand would be required, which again highlights the significant complexity reduction that the channel coefficient methods are able to provide.

The aforementioned measurements and results are repeated with the 20-MHz carriers in Fig. 9, while Fig. 10 plots the corresponding PSDs of the signals at +31 dBm TX power. In these measurements, a purpose-made diode antenna is used as the PIM source. The parameters utilized for these measurements are $M_1 = M_2 = 2$ and $K_1 = K_2 = 3$,

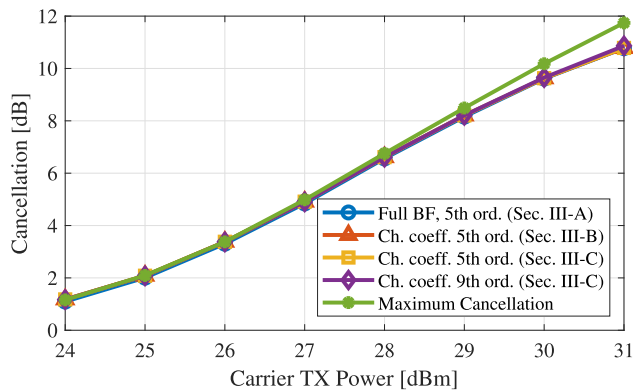


Fig. 9. Power sweep with 20-MHz carriers, diode antenna as PIM source, $M_1 = M_2 = 2$, and $K_1 = K_2 = 3$.

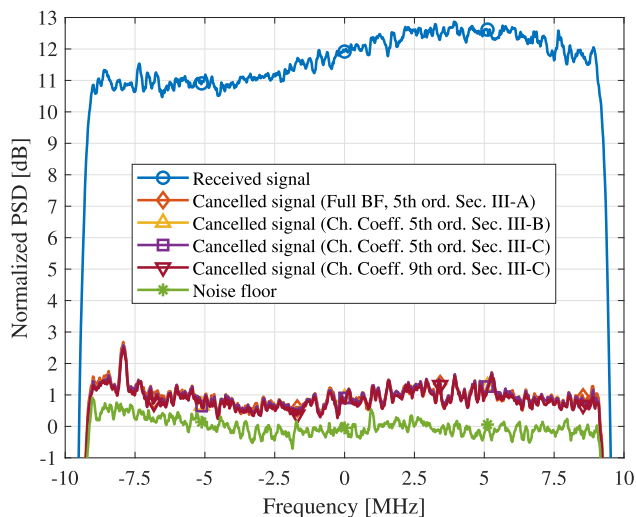


Fig. 10. Cancellation of the 20-MHz carrier PIM from a diode antenna with +31-dBm TX power, $M_1 = M_2 = 2$, and $K_1 = K_2 = 3$.

meaning that the full BF model incorporates now 1880 BFs in $\Phi[t]$, while the channel coefficient models incorporate only 135 BFs. A similar conclusion can be drawn here as with the earlier 5-MHz case, namely, the fifth-order modeling of the channel coefficient models matches extremely well with the performance of the full BF-based fifth-order model, even with increased carrier BW. Again, with high transmit powers, the residual interference power starts increasing, being around 0.9 dB with +31 dBm TX power, where the achieved cancellation is 10.8 dB, while the maximum is 11.7 dB. Unlike in the 5-MHz case, there is no clear advantage in using the ninth-order model in this scenario, as the performance level is on par with the fifth-order models. This is likely due to the nonlinear characteristics of the utilized diode in the PIM source antenna.

In general, it can be observed from both Figs. 7 and 9 that, in order to avoid generating interference to the RX path, the TX power of each carrier needs to be backed off considerably—if no PIM cancellation is carried out. For example, let us allow an interference level of 2 dB above the noise floor in our system, where a single PIM source generates PIM distortion to the RX chain. It can be seen from both Figs. 7 and 9 that employ the cancellation schemes adhere

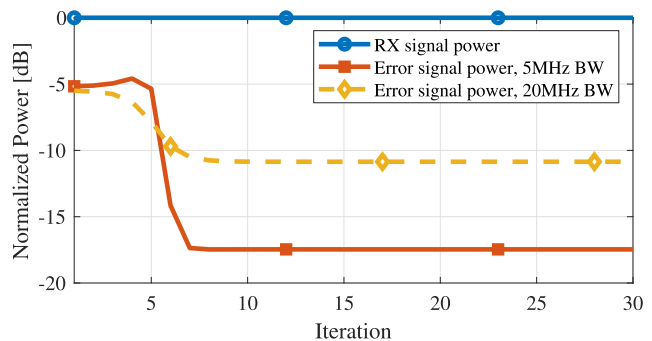


Fig. 11. Convergence of the canceled signal power in the single PIM source cases, with 5- and 20-MHz carriers, +31-dBm TX power, and $N_P = 1$. In the 5-MHz case, $\lambda[0] = 0.2$, and in the 20-MHz case, $\lambda[0] = 2$.

to this limit, even when transmitting at +31 dBm per carrier. On the other hand, if no PIM cancellation is performed, it can be approximated that, in order to induce interference levels of only 2 dB above the noise floor, we would have to set the TX power of each carrier to +22 dBm in the 5-MHz case. This corresponds to a back-off of 9 dB per carrier compared to if a canceller is utilized. Likewise, we would need to back-off the TX power by approximately 7 dB per carrier in the 20-MHz case to induce an interference level of only 2 dB above the noise floor. Thus, by deploying the proposed PIM cancellation methods, the transceiver can increase the TX power by 9 and 7 dB per carrier in the 5- and 20-MHz cases, respectively, compared to the situation without PIM cancellation. This, in turn, increases both the TX power-efficiency and the downlink coverage of the network.

Fig. 11 plots the convergence of the Gauss–Newton algorithm used in the single PM source measurements for both the 5- and 20-MHz cases, utilized by the channel coefficient model of Section III-C. The regularization coefficient $\lambda[t]$ shown in the Gauss–Newton algorithm in (20) is initialized to 0.2 in the 5-MHz case and to 2 in the 20-MHz case. In order to avoid bias in the steady-state operation, the λ coefficient is updated using the following rule:

$$\lambda[t + 1] = 0.91\lambda[t] \quad (25)$$

which will effectively remove the biasing after sufficiently many iterations. It can be seen from Fig. 11 that both the 5- and 20-MHz cases converge to a steady-state operation in about eight iterations.

C. Multiple PIM Sources

In this section, we increase the number of PIM sources in the system to two and three, i.e., $N_P = 2$ and $N_P = 3$. In these measurements, we employ diode antennas as the air-PIM sources, which are placed 3 m from the BS unit, with a 1 m separation between the diode antennas.

Fig. 12 plots the PSDs of the signals of interest when 5-MHz carriers are utilized at +31 dBm TX power with $M_1 = M_2 = 2$ and $K_1 = K_2 = 1$. Here, only fifth-order modeling is employed. First, it is evident that the channel coefficient method from Section III-B cannot handle the coexistence of multiple PIM sources in the system, as was

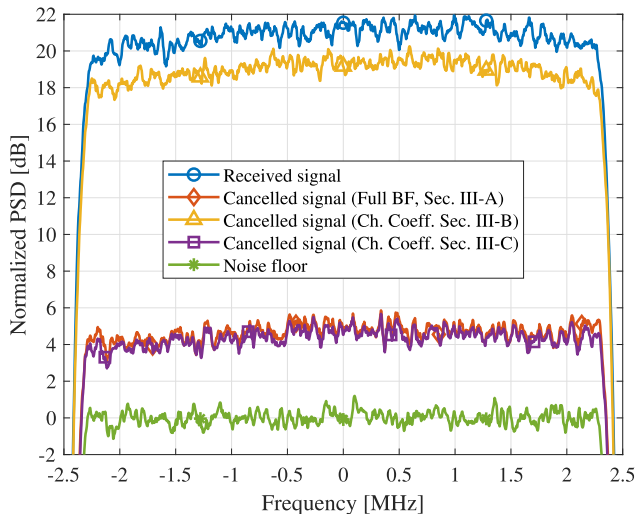


Fig. 12. Cancellation with two diode antennas as PIM sources, 5-MHz carriers, with +31-dBm TX power, $M_1 = M_2 = 2$, and $K_1 = K_2 = 2$.

expected based on the modeling of the system. Therefore, the achievable cancellation of the single PIM source channel coefficient model of Section III-B is merely 2 dB. On the other hand, the fifth-order full BF canceller is capable of mitigating the received PIM by some 16.6 dB, out of the maximum of 20.9 dB when referencing the noise floor. Similar performance is evidenced with the fifth-order channel coefficient model of Section III-C, where N_P is now set to 2. With the chosen parameterization, the full BF model again employs 760 BFs in the matrix $\Phi[t]$, while the multiple PIM source channel coefficient model now employs 130 BFs. Still, with the slightly elevated complexity compared to the single PIM source case, a substantial reduction in complexity is gained by employing the channel coefficient model, without penalty to the cancellation performance.

Fig. 13 plots the PSDs of the signals of interest when 20-MHz carriers are transmitted at +31-dBm TX power. Here, the parameters are set to $M_1 = M_2 = 2$ and $K_1 = K_2 = 3$. As was the case with the 5-MHz carriers, the cancellation performance of the single PIM source channel coefficient model from Section III-B is poor, at only around 2 dB. Meanwhile, both the full BF model and multiple PIM source channel coefficient model from Section III-C achieve cancellation levels of around 13.9 dB, out of a maximum of around 15.2 dB, indicating that the multiple PIM source channel coefficient method is capable of providing the saving in complexity without loss of performance also with higher BWs. In this case, the full BF model incorporates again 1880 BFs, while the multiple PIM source channel coefficient model incorporates 270 BFs.

Next, three PIM antennas are utilized as PIM sources in Fig. 14 when 5-MHz carriers are transmitted. The cancellers follow the same parameterization as above with the 5-MHz carriers: $M_1 = M_2 = 1$ and $K_1 = K_2 = 1$. As before, the single PIM source channel coefficient-based canceller of Section III-B is clearly not capable of handling the PIM stemming from multiple coexisting sources. Meanwhile, the full BF canceller and the multiple PM source channel coefficient-based cancellers are able to suppress the PIM by around 16.5 and 16 dB, respectively, while the maximum cancellation is around

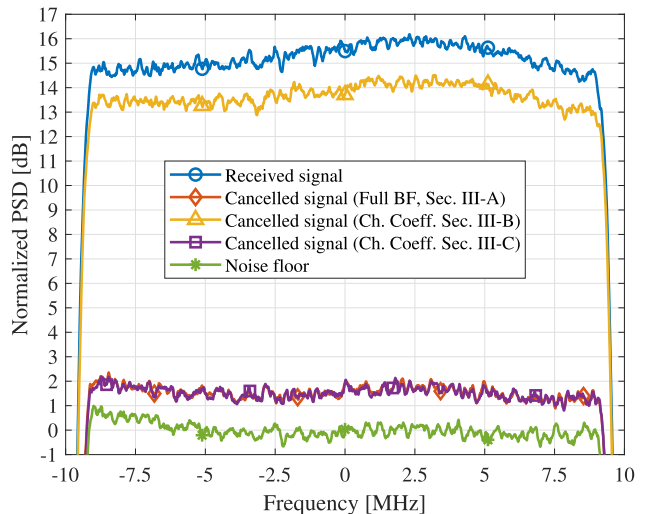


Fig. 13. Cancellation with two diode antennas as PIM sources, 20-MHz carriers, with +31-dBm TX power, $M_1 = M_2 = 2$, and $K_1 = K_2 = 3$.

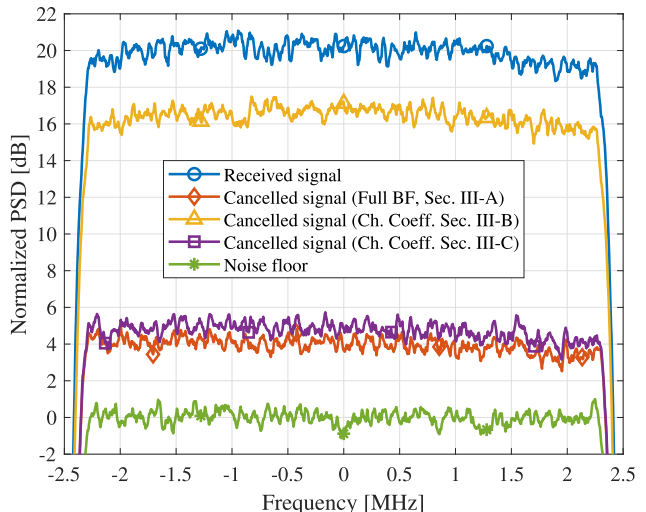


Fig. 14. Cancellation with three diode antennas as PIM sources, 5-MHz carriers, with +31 dBm TX power, $M_1 = M_2 = 2$, and $K_1 = K_2 = 1$.

20 dB. The full BF model again incorporates 760 BFs, while the multiple PIM source channel coefficient-based method employs 195 BFs, which still offers a considerable reduction in complexity in the main path processing.

Finally, we observe and show the convergence of the canceled signal powers in the multiple PIM source scenarios. Fig. 15 illustrates the convergence of the canceled signal powers in the two PIM source cases with 5- and 20-MHz BWs. The parameter vector $\mathbf{a}[t]$ is again initialized to all ones, and the $\lambda[t]$ coefficient is initialized as 0.1 in the 5-MHz case and to 0.15 in the 20-MHz case. The $\lambda[t]$ coefficient is updated using the rule from (25). The initial drop in the error signal power occurs again in around eight iterations, but, to reach full convergence, around ten iterations are required in the 20-MHz case, while the 5-MHz case requires around 13 iterations to achieve convergence. Fig. 16 shows the convergence of the canceled signal power in the three PIM source cases with 5-MHz carriers and with two different initial conditions for $\lambda[0]$, namely, $\lambda[0] = 0.1$ and $\lambda[0] = 0.2$. It can be seen that the initial convergence until

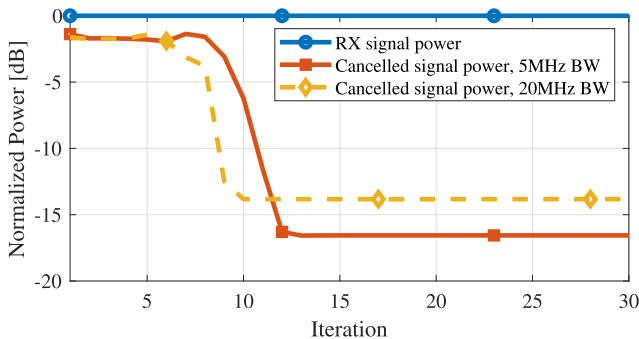


Fig. 15. Convergence of the canceled signal power with two diode antennas as PIM sources with 5- and 20-MHz carriers, +31-dBm TX power, and $N_p = 2$. In the 5-MHz case, $\lambda[0] = 0.1$, and in the 20-MHz case, $\lambda[0] = 0.15$.

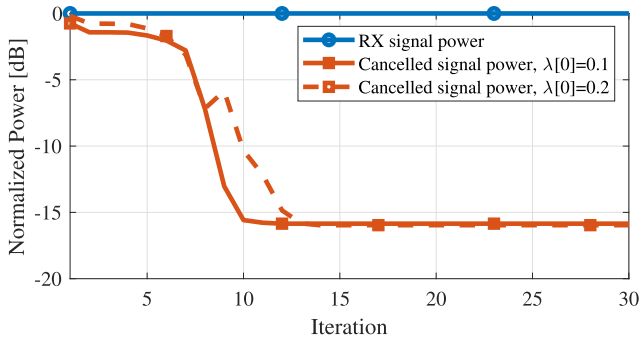


Fig. 16. Convergence of the canceled signal power with three diode antennas as PIM sources with 5-MHz carriers, +31-dBm TX power, and $N_p = 3$, with $\lambda[0] = 0.1$ and $\lambda[0] = 0.2$.

iteration 8 is similar; however, while the power in the $\lambda[0] = 0.1$ case continuously drops until converging at iteration 11, the power in the $\lambda[0] = 0.2$ case has a slight peak starting from iteration 9, which delays the convergence by some iterations. The $\lambda[0] = 0.2$ case then converges at iteration 13.

The convergence plots in Figs. 11, 15, and 16 show that, in general, it is difficult to estimate the exact convergence properties of the Gauss–Newton algorithm with different numbers of PIM sources. This is especially true when the number of PIM sources keeps increasing, as the behavior of the error surface w.r.t. the parameters in the vector $\mathbf{a}[t]$ becomes challenging to predict. This stems from the convergence being dependent on the exact position and orientation of the PIM sources w.r.t. the TX and RX antennas, which inevitably vary between the measurement cases. In Figs. 11, 15, and 16, it can be seen that, in the measured cases, the single PIM source case converges the fastest, which makes intuitive sense. However, the convergence in the two PIM source cases is slightly slower than in the three PIM source case when the same initial value for $\lambda[0]$ is applied. Furthermore, the sensitivity to the initial conditions increases with the number of PIM sources, as evidenced by Fig 16. The slightly erratic performance when $\lambda[0] = 0.2$ could be amended by decreasing λ more slowly, yet that would delay the convergence even further. We omit further analysis of the convergence here, but it presents an interesting possible future work item.

D. Discussion and Future Considerations

Here, we will shortly discuss selected further details, and potential extensions and future work of the introduced PIM

cancellation methods. An important detail of the proposed channel coefficient methods is the assumption that the channel between the transceiver and an individual air-PIM source can be approximated by a single complex tap. As mentioned before, this is a fair assumption in most circumstances and BS deployment scenarios when relatively low BW signals are utilized (such as in the measurements above) since the PIM source requires a line-of-sight channel to the transceiver in order to generate detectable and harmful levels of PIM. Higher system BW could, in general, introduce frequency selectivity and therefore considerable memory effects to the channel model. However, at higher BWs, the PIM issue is typically less severe due to the reflected PIM power being spread over larger BW (lower PIM spectral density), therefore posing less of a threat to the RX sensitivity.

An additional valid note is that our analysis has assumed the PIM sources to be static, i.e., they do not move, and therefore, the channels between the transceiver and the PIM sources are approximately time-invariant. This is a valid assumption in many real-life situations, such as when the air-induced PIM is generated by nuts and/or bolts in the vicinity of the transceiver. However, in certain situations, the PIM may be generated by nonrigid objects that can move or vibrate, such as a steel fence vibrating in the wind, consequently altering the channel between the transceiver and the PIM source. The full BF method and the single PIM source channel coefficient method of Section III-B can inherently handle these changes within one processing block, as the adaptation or parameter estimation is essentially based on block LS in both models. The performance of the multiple PIM source channel coefficient method of Section III-C, on the other hand, is likely to suffer from moving PIM sources, as the solution can only be obtained iteratively. Though this is strictly speaking a matter of the available computational power.

In this work, we have considered the specific case of two carrier frequencies and two spatially multiplexed signals per carrier while canceling the intermodulation products falling on the frequency $2f_1 - f_2$. As mentioned before, the generic modeling results in (5) can be used to determine the BFs with any number of signals per frequency, falling on any odd intermodulation frequency. Furthermore, in systems with more than two carrier frequencies, the pairwise modeling and cancellation of the PIM products, say, $\psi_a[t]$ and $\psi_b[t]$ falling on frequencies of the form $2f_a - f_b$, can be performed as presented in this article; however, there are then also additional third-order intermodulation frequencies. For example, in the case of three carrier frequencies, the instantaneous third-order BFs falling on the frequency $f_a + f_b - f_c$ take the form of $\psi_a[t]\psi_b[t]\psi_c^*[t]$ [4]. In such cases, the presented channel coefficient methods are not directly valid in their current form but would require slight reformulations.

Finally, the PIM sources may also generate harmonic distortion, which is neglected in the system modeling in this work. In general, the harmonic products appear far from the passband in the frequency domain since the second harmonic frequencies are of the form $f_a + f_b$. Yet, there can be cases where, say, band n3 and n78 BSs are colocated, and in such cases, harmonic PIM generated from the band n3 BS can

TABLE IV
INSTANTANEOUS ($m = 0$) BFs OF THE FULL BF METHOD UP TO POLYNOMIAL ORDER $p = 5$

p	Unique Instantaneous BFs	Number of BFs	
		$k = 0$	$k \neq 0$
3	$x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]$, $x_{1,1}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]$, $x_{1,1}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]$, $x_{1,2}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]$, $x_{1,2}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]$	6	8
5	$x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,1}[t]x_{1,1}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,1}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,1}^*[t-k]x_{2,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{1,2}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,1}[t-k]^2$, $x_{1,2}[t]x_{1,2}[t-k]x_{2,2}^*[t-k]x_{2,2}[t-k]^2$	34	48

reflect the band n78 BS and desensitize the corresponding RX path. In such cases, new second-order BFs are required in the form of $\psi_a[t]\psi_b[t]$ in order to cancel such harmonic PIM.

V. CONCLUSION

In this article, we have investigated air-induced PIM and its efficient digital cancellation in a challenging FDD MIMO scenario. As a fundamental basis, extensive system modeling was provided, describing the total PIM signal at a given intermodulation frequency, when an arbitrary number of transceivers operate simultaneously and when there are an arbitrary number of PIM sources causing distortion. Stemming from such a system model, a corresponding digital cancellation scheme based on the full set of polynomial BFs, called the full BF method, was then described. To alleviate the high computational complexity of the full BF approach, an alternative method leveraging the air-PIM channel coefficients, called the channel coefficient method, was formulated. While the basic channel coefficient method allows only for handling the distortion stemming from a single dominant PIM source, we then proposed a novel Gauss–Newton-based iterative method, which can identify the channel coefficients of any number of concurrent PIM sources. The provided computational complexity analysis of all three methods reveals substantial complexity reduction when employing the proposed channel coefficient-based solutions compared to the full BF model—particularly when the number of significant PIM sources is moderate. All the methods were assessed and validated through extensive RF measurements at the 5G NR band n3, utilizing true BS hardware. The measurement results show that the high cancellation fidelity of the full BF model can be retained with much reduced complexity with the proposed channel coefficient-based methods. Our future work topics include addressing the potential time-varying nature of the air-PIM channels in realistic BS deployment sites while also generalizing the PIM cancellation solutions to handle harmonic PIM.

APPENDIX I BASIS FUNCTIONS OF THE FULL BF MODEL

For readers’ convenience, the instantaneous BFs of the full BF model are given in Table IV up to the polynomial order of 5. These are stemming from (6).

APPENDIX II PARTIAL DERIVATIVES OF THE FUNCTION $\hat{r}[t]$

The partial derivatives of the function $\hat{r}[t]$, defined in (18), w.r.t. the parameters $a_{s,1,1}$, $a_{s,1,2}$, $a_{s,2,1}$, $a_{s,2,2}$, and $\gamma_{s,0,1,0}$ are given as

$$\frac{\partial \hat{r}[t]}{\partial a_{s,1,1}} = 2\gamma_{s,0,1,0}(a_{s,1,1}a_{s,2,1}^*x_{1,1}^2[t]x_{2,1}^*[t] + a_{s,1,2}a_{s,2,1}^*x_{1,1}[t]x_{1,2}[t]x_{2,1}^*[t] + a_{s,1,1}a_{s,2,2}^*x_{1,1}^2[t]x_{2,2}^*[t] + a_{s,1,2}a_{s,2,2}^*x_{1,1}[t]x_{1,2}[t]x_{2,2}^*[t]) \quad (26)$$

$$\frac{\partial \hat{r}[t]}{\partial a_{s,1,2}} = 2\gamma_{s,0,1,0}(a_{s,1,1}a_{s,2,1}^*x_{1,1}[t]x_{1,2}[t]x_{2,1}^*[t] + a_{s,1,2}a_{s,2,1}^*x_{1,2}^2[t]x_{2,1}^*[t] + a_{s,1,1}a_{s,2,2}^*x_{1,1}[t]x_{1,2}[t]x_{2,2}^*[t] + a_{s,1,2}a_{s,2,2}^*x_{1,2}^2[t]x_{2,2}^*[t]) \quad (27)$$

$$\frac{\partial \hat{r}[t]}{\partial a_{s,2,1}^*} = \gamma_{s,0,1,0}(a_{s,1,1}^2x_{1,1}^2[t]x_{2,1}^*[t] + 2a_{s,1,1}a_{s,1,2}x_{1,1}[t]x_{1,2}[t]x_{2,1}^*[t] + a_{s,1,2}^2x_{1,2}^2[t]x_{2,1}^*[t]) \quad (28)$$

$$\frac{\partial \hat{r}[t]}{\partial a_{s,2,2}^*} = \gamma_{s,0,1,0}(a_{s,1,1}^2x_{1,1}^2[t]x_{2,2}^*[t] + 2a_{s,1,1}a_{s,1,2}x_{n,1,1}[t]x_{1,2}[t]x_{2,2}^*[t] + a_{s,1,2}^2x_{1,2}^2[t]x_{2,2}^*[t]) \quad (29)$$

and

$$\begin{aligned} \frac{\partial \hat{f}[t]}{\partial \gamma_{s,0,1,0}} = & a_{s,1,1}^2 a_{s,2,1}^* x_{1,1}^2[t] x_{2,1}^*[t] \\ & + 2a_{s,1,1} a_{s,1,2} a_{s,2,1}^* x_{1,1}[t] x_{1,2}[t] x_{2,1}^*[t] \\ & + a_{s,1,2}^2 a_{s,2,1}^* x_{1,2}^2[t] x_{2,1}^*[t] \\ & + a_{s,1,1}^2 a_{s,2,2}^* x_{1,1}^2[t] x_{2,2}^*[t] \\ & + 2a_{s,1,1} a_{s,1,2} a_{s,2,2}^* x_{1,1}[t] x_{1,2}[t] x_{2,2}^*[t] \\ & + a_{s,1,2}^2 a_{s,2,2}^* x_{1,2}^2[t] x_{2,2}^*[t]. \end{aligned} \quad (30)$$

REFERENCES

- [1] H. Holma, A. Toskala, and T. Nakamura, Eds., *5G Technology: 3GPP New Radio*, 1st ed. Hoboken, NJ, USA: Wiley, 2019.
- [2] *NR; Base Station (BS) Radio Transmission and Reception*, Standard 3GPP Tech. Spec. 38.104 v17.6.0, (Release 17), Jun. 2022.
- [3] A. Kiayani, M. Abdelaziz, L. Anttila, V. Lehtinen, and M. Valkama, "Digital mitigation of transmitter-induced receiver desensitization in carrier aggregation FDD transceivers," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 11, pp. 3608–3623, Dec. 2015.
- [4] C. Yu, W. Cao, Y. Guo, and A. Zhu, "Digital compensation for transmitter leakage in non-contiguous carrier aggregation applications with FPGA implementation," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 12, pp. 4306–4318, Dec. 2015.
- [5] J. Henrie, A. Christianson, and W. J. Chappell, "Cancellation of passive intermodulation distortion in microwave networks," in *Proc. 38th Eur. Microw. Conf.*, Oct. 2008, pp. 1153–1156.
- [6] J. Henrie, A. Christianson, and W. J. Chappell, "Engineered passive nonlinearities for broadband passive intermodulation distortion mitigation," *IEEE Microw. Wireless Compon. Lett.*, vol. 19, no. 10, pp. 614–616, Oct. 2009.
- [7] Q. Jin, J. Gao, H. Huang, and L. Bi, "Mitigation methods for passive intermodulation distortion in circuit systems using signal compensation," *IEEE Microw. Wireless Compon. Lett.*, vol. 30, no. 2, pp. 205–208, Feb. 2020.
- [8] X. Chen, T. Ren, D. J. Pommerenke, and M. Yu, "Broadband mechanical intermodulation tuner using reconfigurable distributed nonlinearity," *IEEE Trans. Microw. Theory Techn.*, vol. 70, no. 1, pp. 5–13, Jan. 2022.
- [9] H.-T. Dabag, H. Gheidi, P. Gudem, and P. M. Asbeck, "All-digital cancellation technique to mitigate self-jamming in uplink carrier aggregation in cellular handsets," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2013, pp. 1–3.
- [10] H.-T. Dabag, H. Gheidi, S. Farsi, P. Gudem, and P. M. Asbeck, "All-digital cancellation technique to mitigate receiver desensitization in uplink carrier aggregation in cellular handsets," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 12, pp. 4754–4765, Dec. 2013.
- [11] H. Gheidi, H.-T. Dabag, Y. Liu, P. M. Asbeck, and P. Gudem, "Digital cancellation technique to mitigate receiver desensitization in cellular handsets operating in carrier aggregation mode with multiple uplinks and multiple downlinks," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2015, pp. 221–224.
- [12] M. Z. Waheed et al., "Passive intermodulation in simultaneous transmit-receive systems: Modeling and digital cancellation methods," *IEEE Trans. Microw. Theory Techn.*, vol. 68, no. 9, pp. 3633–3652, Sep. 2020.
- [13] J. R. Wilkerson, I. M. Kilgore, K. G. Gard, and M. B. Steer, "Passive intermodulation distortion in antennas," *IEEE Trans. Antennas Propag.*, vol. 63, no. 2, pp. 474–482, Feb. 2015.
- [14] P. L. Lui, "Passive intermodulation interference in communication systems," *Electron. Commun. Eng. J.*, vol. 2, no. 3, pp. 109–118, Jun. 1990.
- [15] Q. Jin, J. Gao, G. T. Flowers, Y. Wu, and G. Xie, "Modeling of passive intermodulation with electrical contacts in coaxial connectors," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 9, pp. 4007–4016, Sep. 2018.
- [16] V. Lampu, L. Anttila, M. Turunen, M. Fleischer, J. Hellmann, and M. Valkama, "Air-induced passive intermodulation in FDD networks: Modeling, cancellation and measurements," in *Proc. 55th Asilomar Conf. Signals, Syst., Comput.*, Oct. 2021, pp. 983–988.
- [17] V. Lampu, L. Anttila, M. Turunen, M. Fleischer, J. Hellmann, and M. Valkama, "Air-induced PIM cancellation in FDD MIMO transceivers," *IEEE Microw. Wireless Compon. Lett.*, vol. 32, no. 6, pp. 780–783, Jun. 2022.
- [18] M. Z. Waheed, D. Korpi, A. Kiayani, L. Anttila, and M. Valkama, "Digital cancellation of passive intermodulation: Method, complexity and measurements," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Aug. 2019, pp. 1–3.
- [19] J. J. Henrie, A. J. Christianson, and W. J. Chappell, "Linear–nonlinear interaction and passive intermodulation distortion," *IEEE Trans. Microw. Theory Techn.*, vol. 58, no. 5, pp. 1230–1237, May 2010.
- [20] G. Macchiarella, G. B. Stracca, and L. Miglioli, "Experimental study of passive intermodulation in coaxial cavities for cellular base stations duplexers," in *Proc. 34th Eur. Microw. Conf.*, vol. 2, Oct. 2004, pp. 981–984.
- [21] A. Shitov, A. G. Schuchinsky, M. B. Steer, and J. M. Wetherington, "Characterisation of nonlinear distortion and intermodulation in passive devices and antennas," in *Proc. 8th Eur. Conf. Antennas Propag. (EuCAP)*, Apr. 2014, pp. 1454–1458.
- [22] E. A. Keehr and A. Hajimiri, "Successive regeneration and adaptive cancellation of higher order intermodulation products in RF receivers," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 5, pp. 1379–1396, Mar. 2011.
- [23] D. Kozlov, S. Bulja, R. Lundy, P. Rulikowski, and J. J. Walsh, "Mitigation of passive intermodulation on planar microstrip circuits with distributed current-driven nonlinearities," in *Proc. 48th Eur. Microw. Conf. (EuMC)*, Sep. 2018, pp. 1029–1032.
- [24] D. Kozlov et al., "Practical mitigation of passive intermodulation in microstrip circuits," *IEEE Trans. Electromagn. Compat.*, vol. 62, no. 1, pp. 163–172, Feb. 2020.
- [25] M. Omer, R. Rimini, P. Heidmann, and J. S. Kenney, "A compensation scheme to allow full duplex operation in the presence of highly nonlinear microwave components for 4G systems," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Jun. 2011, pp. 1–4.
- [26] W. Cao, Y. Li, and A. Zhu, "Digital suppression of transmitter leakage in FDD RF transceivers: Aliasing elimination and model selection," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 3, pp. 1500–1511, Mar. 2018.
- [27] A. Gebhard et al., "A robust nonlinear RLS type adaptive filter for second-order-intermodulation distortion cancellation in FDD LTE and 5G direct conversion transceivers," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 5, pp. 1946–1961, May 2019.
- [28] T. Paireder, C. Motz, and M. Huemer, "Spline-based adaptive cancellation of even-order intermodulation distortions in LTE–A/5G RF transceivers," *IEEE Trans. Veh. Technol.*, vol. 70, no. 6, pp. 5817–5832, Jun. 2021.
- [29] Z. Wang, M. Ma, and F. Qin, "Neural-network-based nonlinear self-interference cancellation scheme for mobile stations with dual-connectivity," *IEEE Access*, vol. 9, pp. 53566–53575, 2021.
- [30] C. Motz, T. Paireder, and M. Huemer, "Low-complex digital cancellation of transmitter harmonics in LTE–A/5G transceivers," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 948–963, 2021.
- [31] M. Abdelaziz, L. Anttila, C. Tarver, K. Li, J. R. Cavallaro, and M. Valkama, "Low-complexity subband digital predistortion for spurious emission suppression in noncontiguous spectrum access," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 11, pp. 3501–3517, Nov. 2016.
- [32] A. Abdelhafiz, A. Kwan, M. Younes, O. Hammi, N. Boulejfen, and F. M. Ghannouchi, "Augmented dual-band digital predistorter for reducing cross-band intermodulation distortion using predictive injection technique," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 11, pp. 3518–3527, Nov. 2016.
- [33] A. Sabharwal, P. Schniter, D. Guo, D. W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1637–1652, Sep. 2014.
- [34] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 7, pp. 3025–3041, Jul. 2019.
- [35] K. E. Kolodziej, *In-Band Full-Duplex Wireless Systems Handbook*. Norwood, MA, USA: Artech House, 2021.
- [36] J.-B. Yan, S. Yong, and J. T. Bernhard, "Intermodulation and harmonic distortion in frequency reconfigurable slot antenna pairs," *IEEE Trans. Antennas Propag.*, vol. 62, no. 3, pp. 1138–1146, Mar. 2014.
- [37] S. Akoum and J. Acharya, "Full-dimensional MIMO for future cellular networks," in *Proc. IEEE Radio Wireless Symp. (RWS)*, Jan. 2014, pp. 1–3.
- [38] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

- [39] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.
- [40] X.-S. Yang, *Optimization Techniques and Applications With Examples*. Hoboken, NJ, USA: Wiley, 2018.
- [41] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1996.



Vesa Lampu (Graduate Student Member, IEEE) received the M.Sc. degree in electrical engineering from Tampere University, Tampere, Finland, in 2019, where he is currently pursuing the D.Sc. degree at the Department of Electrical Engineering.

He has been working at the Department of Electrical Engineering, Tampere University, since 2018. His research focuses on energy-efficient multiple-input–multiple-output (MIMO) systems. He is also involved in research on nonlinear system identification, with applications in full-duplex, digital predistortion techniques, and passive intermodulation cancellation.

tortion techniques, and passive intermodulation cancellation.



Lauri Anttila received the D.Sc. (Tech.) degree (Hons.) from the Tampere University of Technology (TUT), Tampere, Finland, in 2011.

Since 2016, he has been a University Researcher with the Electrical Engineering Unit, Tampere University, Tampere. From 2016 to 2017, he was a Visiting Research Fellow with Aalto University, Espoo, Finland. His research interests are in radio communications, radio-based sensing, and signal processing, with emphasis on the radio implementation challenges in 5G/6G mobile networks, large-scale antenna systems, and full-duplex radio.

scale antenna systems, and full-duplex radio.



Matias Turunen is currently a Researcher and a Laboratory Specialist with the Department of Electrical Engineering, Tampere University, Tampere, Finland.

His research interests include inband full-duplex radios with an emphasis on analog RF cancellation, OFDM radar, and 5G new radio (NR) systems.



Marko Fleischer received the Diploma degree in electrical engineering (EE) from the Technical University of Dresden (TU Dresden), Dresden, Germany, in 1996.

In 1990, he worked in the defense industry. He joined Siemens Mobile Communication, Munich, Germany, in 2000. Since 2000, he has been involved in the development of new radio (NR) concepts ranging from active to passive intermodulation cancellation methods [active intermodulation (AIM)/passive intermodulation (PIM)/self-interference cancellation (SIC)], studies of 5G full duplex, and OFDM radar. He is currently with Nokia RF System Engineering, Ulm, Germany.

He is currently with Nokia RF System Engineering, Ulm, Germany.



Jan Hellmann received the Diploma degree in electrical and computer engineering from the Technical University of Dresden, Dresden, Germany, in 1990.

With 30 years of professional experience, he held various functions in the field of research and development for mobile radio communications at Siemens, Munich, Germany, and Nokia, Ulm, Germany. He is currently working at Nokia RF System Engineering, Ulm. His interests are in RF-related signal processing algorithms, such as passive intermodulation cancellation.



Mikko Valkama (Fellow, IEEE) received the D.Sc. (Tech.) degree (Hons.) in electrical engineering (EE) from the former Tampere University of Technology (TUT), Tampere, Finland, in 2001.

In 2003, he was working as a Visiting Post-Doctoral Research Fellow at the Communications Systems and Signal Processing Institute, San Diego State University (SDSU), San Diego, CA, USA. He is currently a Full Professor and the Head of the Electrical Engineering Unit at the newly formed Tampere University, Tampere.

His research interests include radio communications, radio localization, and radio-based sensing, with an emphasis on 5G and 6G mobile networks.