Tight Upper Bound Performance of Full-Duplex MIMO-BICM-IDD Systems in the Presence of Residual Self-Interference

Mohamad A. Ahmed, *Member, IEEE*, and Charalampos C. Tsimenidis, *Senior Member, IEEE*

***Abstract*— In this paper, we derive a tight upper bound on the performance of a coded full-duplex multiple-input multiple-output (MIMO)-based bidirectional transceiver. Iterative detection and decoding (IDD) are proposed to suppress the residual self-interference (SI) remaining after applying different stages of SI cancellation. IDD comprises an adaptive minimum mean-squared error filter with log-likelihood ratio demapping, while the soft decoder by using soft-in soft-out decoding utilizes the maximum *a posteriori* algorithm. Furthermore, bit-interleaved coded modulation is considered in the presence of additive white Gaussian noise over MIMO frequency non-selective Rayleigh fading channels. Simulation results are presented to demonstrate the bit-error rate (BER) performance as a function of the signal-to-noise ratio showing a close match to the SI-free case for the proposed system. Furthermore, we validate our results by deriving a tight upper bound on the performance of the proposed system using rate-1/2 convolutional codes together with M-ary quadrature amplitude modulation, which asymptotically exhibits a close agreement with the simulated BER performance. Moreover, extrinsic information transfer chart analysis is used to investigate the convergence behavior of the proposed IDD receiver and to determine the number of iterations required for this convergence.**

***Index Terms*— Full-duplex (FD), multiple-input multiple-output (MIMO), self-interference cancellation (SIC), iterative detection and decoding (IDD), performance bounds, and EXIT chart.**

I. Introduction

FULL-DUPLEX multiple-input multiple-output (FD-MIMO) wireless communication has become one of the most promising techniques that enhance spectral efficiency, channel capacity and diversity gain. This is due to the fact that FD-MIMO transceivers always send and receive data simultaneously using the same frequency via multiple antennas. A key challenge in FD-based wireless communications is self-interference (SI) which degrades significantly system performance. SI is caused principally by the signals transmitted from the FD transceiver, which exhibit larger energy than the desired incoming signals due to path loss propagation phenomena. The large power difference between the signal of interest, which arrives from a distant source and needs to be detected, and the SI signal transmitted by the FD transceiver itself poses an extreme challenge to the receiver designer [1], [2].

Thus far, several methods have been proposed to overcome SI, which can be grouped into two categories, i.e. passive and active self-interference cancellation (SIC). Passive methods rely on the physical separation of the transmit and receive antennas in order to increase the isolation loss between them, and hence, reduce the magnitude of local interference. This approach is called antenna separation. In contrast, active approaches are implemented either in the analogue domain, which is always before the analog-to-digital converter (ADC), or in the digital domain, and thus after the ADC. Precise knowledge of the interfering signal and its channel can be utilized to subtract it either in the bandpass or complex baseband. Additionally, SIC employing spatial domain mitigation, based on zero-forcing and null-space projection, can be utilized for this purpose [3]–[10]. Furthermore, a combination of different domain methods can be used to obtain improved performance [11]–[13].

Iterative soft parallel interference cancellation (soft-PIC) using the soft decoding via soft-in soft-out iterative detection and decoding (IDD) can be exploited with the FD-MIMO transceiver for SIC [8]. This approach can be implemented by employing channel encoding using convolutional or turbo codes followed by an interleaver to combat burst errors. The coded and interleaved bits are then mapped into symbols according to the modulation scheme utilized. These approach is referred to as bit-interleaved coded modulation (BICM) [14].

Additionally, orthogonal frequency division multiplexing (OFDM) can be employed in conjunction with MIMO-BICM to enhance spectral efficiency and combat the effect of frequency-selective fading channels [15]–[18]. The concept of using OFDM to convert a frequency selective multipath channel into a set of parallel frequency-flat fading channels is well-known and widely used in the literature. This assumption is valid for OFDM as long as the cyclic prefix (CP) used has sufficient length to cover the delay spread of the multipath channel [3] and the references within. Therefore, the results

Manuscript received December 28, 2016; revised May 29, 2017 and August 27, 2017; accepted October 24, 2017. Date of publication November 3, 2017; date of current version January 8, 2018. This work was supported in part by the Engineering and Physical Sciences Research Council under Project EP/N004299/1 and in part by the Cooperative Backhaul Aided Next-Generation Digital Subscriber Loops. The associate editor coordinating the review of this paper and approving it for publication was O. O. Koyluoglu. *(Corresponding author: Mohamad A. Ahmed.)*

M. A. Ahmed is with the College of Electronics Engineering, Ninevah University, Mosul 41002, Iraq (e-mail: mohamad.alhabbar@uoninevah.edu.iq).

C. C. Tsimenidis is with the School of Electrical and Electronic Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K. (e-mail: charalampos.tsimenidis@newcastle.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2017.2768027

presented in this paper are valid for appropriately designed OFDM systems, where the effects related to frequency-offset induced inter-carrier interference have been compensated.

Iterative detection and decoding (IDD) techniques, which are commonly referred to as turbo-like receivers or turbo-BLAST (BLAST stands for Bell Labs layered space time), are implemented in the MIMO receiver using several iterative exchanges of soft information between the detector and decoder [8], [10], [19], [20]. IDD offers high throughput, efficiency, and reliable performance in the case of noise and interference impaired environments in wireless communications [8], [14], [21], [22]. Soft information may take the forms of maximum *a posterior* probability (MAP), maximum likelihood (ML), or a log-likelihood ratio (LLR) [8], [16]. Such systems have been proposed and successfully employed for the mitigation of various types of wireless interference, such as inter-symbol interference (ISI), co-channel interference (CCI), co-antenna interference (CAI), multiple-access (multi-antenna) interference (MAI), and multi-user interference (MUI) [19], [23]–[25]. In [10], a tight upper bound on the performance of the coded FD single-input multiple-output (SIMO) was derived in the presence of residual SI after applying passive and active SIC.

In practice, the FD-MIMO system using IDD can be considered as a traditional two-user MIMO system, where there is a large power difference between the users. In such scenario, challenges arise to deal with the dynamic range of the ADCs and the effects of the residual SI [3], [26]. The local availability of the perfect SI symbols allows different cancellation approaches for the SI, which would not have been possible if the system was treated as a traditional MIMO system. There are also challenges associated with the estimation of the SI related channels that might need to be calculated in HD mode prior to the FD operation [12], [27].

Theoretical and experimental studies of the exploitation of BLAST using the turbo principles for MIMO in [19] have shown that this combination gives higher data rates, better performance and more efficient computation compared to conventional BLAST schemes with a maximum likelihood detector under the same conditions and in Rayleigh-fading scenarios [20].

An improved IDD architecture with anti-gray mapping has been proposed to overcome the slow convergence rate when employing gray mapping with the traditional turbo-BLAST in a MIMO wireless communication system [21]. This improved scheme was achieved by applying the *posteriori* LLR information from the outer of the soft decoder to the detector, rather than the *extrinsic* information, which is always used in conventional IDD, in order to produce faster convergence. A similar idea has been used with OFDM in [17] for a comparison between the linear and non-linear minimum mean-squared error (MMSE) detectors. The first detector produces *a priori* LLR information as a result of processing and equalizing the received symbols and the extrinsic soft information as the traditional turbo-BLAST, while the non-linear MMSE detector replaces the extrinsic information with *posteriori* LLR information, which is fed directly back from the soft decoder.

In another study [22], the authors dealt with the reduction of the high computational complexity of IDD based on soft interference cancellation (soft-IC) by simplifying the instantaneous matrix inversions at the cost of a slight reduction in performance. A modified turbo-BLAST with OFDM has been considered in [15] based on soft-PIC and MMSE filtering for the MIMO frequency selective fading channel and in the presence of imperfect channel estimation. The researchers then proposed hard-decision based equalization using bit-level interference cancellation as a form of low complexity IDD at the cost of a slight degradation in the overall system performance [16].

The systems presented in [24] and [25] exploited IDD using soft-IC that can be achieved by firstly creating a soft copy of interfering symbols using the LLR information supplied by the outer soft decoder to the soft-IC/MMSE detector. These estimated symbols can then be combined with the interfering channel, which may be known or need to be estimated. Finally, a subtraction from the received signal is implemented in order to obtain the desired signal, however, a linear adaptive filter is required to cancel the residual interference.

The systems in [28] and [29] consider low-density parity check (LDPC) codes as another type of powerful linear error correcting block codes, which can be used to achieve reliable transmission of data that can approach the capacity bounds. In the system proposed in [28], an iterative detection and decoding is assumed for transmission using single antenna in each terminal and by utilizing hybrid relaying of half-duplex and out-of-the-band FD. The latter uses simultaneous transmitting and receiving using different frequency bands, which consequently does not produce SI. Moreover, the system in [29] assumed bidirectional in-band FD-MIMO communication, in which the constellation shaping and forward error correction are exploited to mitigate the residual SI which is modelled as Rician distribution with small κ-factor. However, it was reported that the performance of LDPC codes is degraded and the complexity of the decoding algorithms is increased when low code rates of transmission and/or short lengths of frames are used for transmission [30]. Due the above reasons, LDPC was not considered in our paper.

The key contributions of this paper are to extend our previous works in [8] and [10], in which simulation results were presented for FD-MIMO-IDD in [8], and quadrature phase shift keying (QPSK) modulation scheme is employed without any performance analyses. While in [10], FD-SIMO-IDD is used for QPSK modulation scheme with a tight upper bound on the performance of the proposed system based on SIMO transmission only. In this paper, IDD is exploited in the context of coded FD-MIMO with M-ary quadrature amplitude modulation (QAM), to mitigate more effectively the residual SI remaining after assuming of multiple stages of SIC are applied. After a number of iterations, the performance achieved is shown to be very close to the SI-free case. Additionally, in order to validate the simulation results, we derive a tight upper bound on the performance of the proposed FD-MIMO for rate-1/2 convolutional codes, which shows asymptomatic close agreement with the simulated performance, and this tight bound can be used for performance

evaluation of such a system without the need to implement time-consuming simulations. Moreover, extrinsic information transfer (EXIT) chart analysis is employed as a semi-analytic method to visualize the convergence behavior between the proposed detector and decoder, and to evaluate the number of iterations required for this convergence.

The rest of the paper is organized as follows. In Section II, the signal and system models are introduced. Firstly, we describe the overall coded FD-MIMO bi-directional transceiver by defining the channel and noise characteristics. Moreover, the equations for the transmitted and received signals are given in Subsections II-A and II-B. Subsection II-C is to discuss the channel estimation techniques employed for FD-MIMO transceivers. In Subsection II-D, we present different stages of SIC in the passive and active domains. Additionally, our proposed IDD system is presented in Subsection II-E to combat the residual SI by utilizing iterative soft-PIC with adaptive MMSE filtering for coded FD-MIMO. In Section III, we derive a tight upper bound on the performance of the proposed IDD system in the presence of residual SI. Moreover, in Section IV, we employ EXIT chart analysis as a semi-analytic technique to analyse and visualize the convergence behavior of the proposed iterative system. Section V presents simulation results and discussion, and finally, the paper's conclusions are drawn in Section VI.

Notation: Matrices and vectors are denoted by uppercase and lowercase boldface characters, respectively. The transpose, conjugate and Hermitian of matrix $\mathbf{A}$ are denoted by $\mathbf{A}^T$, $\mathbf{A}^*$ and $\mathbf{A}^{\mathcal{H}}$, respectively. $\|\mathbf{a}\|$, $\|\mathbf{A}\|$, are the Euclidean norms of the vector $\mathbf{a}$ and the matrix $\mathbf{A}$, respectively, while, $|\mathbf{A}|$ is the determinant of the matrix $\mathbf{A}$. Also $\Gamma(a)$ is the complete Gamma function of variable a. $\mathbb{E}[\mathbf{x}]$ is the statistical expectation of the random vector $\mathbf{x}$. The binomial coefficient is defined as $\binom{a}{b} = \frac{a!}{b!(a-b)!}$ where $n!$ is the factorial operation for a positive integer number n. Finally, $\Re[A]$ and $\Im[A]$ represent for the real and imaginary parts of A, respectively.

II. Signal and System Model

In this paper, a wireless scenario is considered in which two bi-directional point-to-point nodes a and b are communicating using FD-MIMO transceivers [8], [10], [12]. We assume that each node is equipped with N_{tx} transmit antennas and N_{rx} receive antennas, where $N_{tx} \leq N_{rx}$ as shown in Fig. 1, for the sake of increasing the diversity gain and also to reduce the source of SI and transmit noise [3]. The four wireless links illustrated in this figure are denoted as $\mathbf{H}_{ab}$ and $\mathbf{H}_{ba}$ indicating the forward channels, while $\mathbf{H}_{aa}$ and $\mathbf{H}_{bb}$ denote the self interference channels. All channels in this paper, except the SI channel, are considered to be frequency non-selective Rayleigh fading channels, i.e. frequency-flat fading channels, which can be defined generally as $\mathbf{H} \sim \mathcal{CN}(0, \mathbf{I}_{N_{tx}N_{rx}})$ [3], [26], [31], [32]. In contrast, the SI channel, $\mathbf{H}_{bb}$ is modelled as a Rician distribution, i.e. $\mathbf{H}_{bb} \sim \mathcal{CN}(\kappa, \mathbf{I}_{N_{tx}N_{rx}})$ for node b, with small κ-factor resulting from the imperfect cancellation of the LoS propagation path [33] that approaches zero if multiple stages of SIC are utilized, in which case can be modelled using a Rayleigh distribution [32].

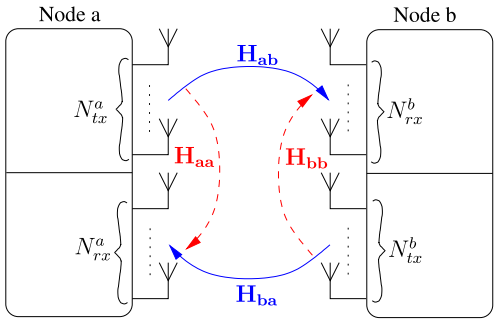

Fig. 1. Bi-directional Full-duplex MIMO Transceiver.

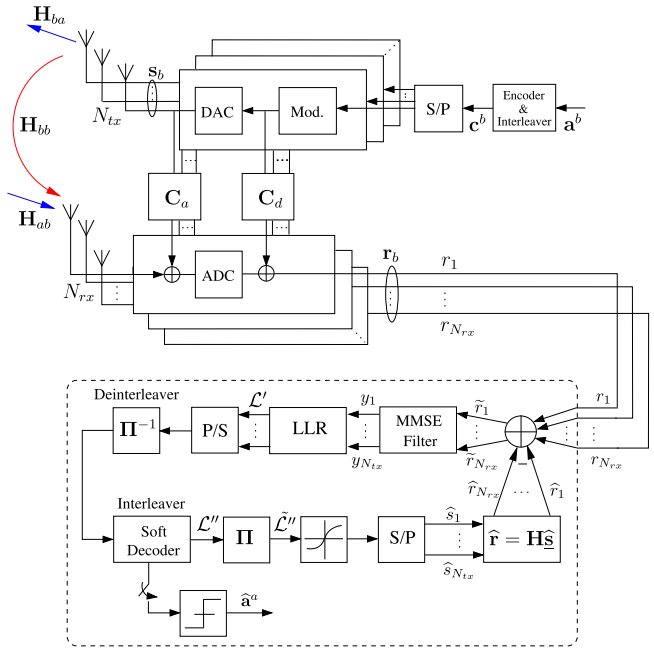

Fig. 2. Bi-directional FD-MIMO Transceiver at node b.

Moreover, the additive white Gaussian noise (AWGN) samples at the receive port of any node are independent and identically distributed (i.i.d.) with zero mean and variance $\sigma^2_{v_i}$, and can be defined as $\boldsymbol{v}_i \sim \mathcal{CN}(0, \sigma^2_{v_i}\mathbf{I}_{N_{rx}})$, and $\sigma^2_{v_i} = \mathbb{E}[\boldsymbol{v}_i \boldsymbol{v}_i^{\mathcal{H}}]$, where $i \in \{a, b\}$.

A. Transmitter Structure

Without loss of generalization, this paper focuses not only on the link from node a to node b, but also on mitigating SI caused by node b's transmitter to its receive port. At the transmitter of node b, as shown in Fig. 2, the q^{th} bit, a_q^b, in a sequence of information bits, $\mathbf{a}^b$ with total length of Q, is passed through a recursive systematic convolutional (RSC), then interleaved to create the p^{th} encoded and interleaved bit, c_p^b, which belongs to the bits sequence, $\mathbf{c}^b$, of total length P. The code rate is defined as $R_c \in [0, 1] = Q/P$.

Subsequently, the coded bits are interleaved using a random interleaver, $\Pi(.)$, and the interleaved bits are divided using a serial-to-parallel multiplexer into N_{tx} sub-streams. In each branch, the conversion of bits to symbols, i.e. mapping,

is achieved according to the modulation scheme used. Each vector in the n^{th} antenna, $\mathbf{c}_n \triangleq [c_{n,1}, \ldots, c_{n,m}] \in \{0,1\}^m$, where $n = 1, 2, \ldots, N_{tx}$, and m represents the number of bits per each symbol, is mapped according to the function

$$\begin{aligned} &\psi(\mathbf{c}_n) : \{0,1\}^m \rightarrow \mathcal{S}, \\ &s_n = \psi(\mathbf{c}_n) \in \mathcal{S}, \end{aligned} \tag{1}$$

where $\mathcal{S} = \{\alpha_1, \alpha_2, \ldots, \alpha_M\}$, α_τ represents a set of all possible complex symbols, $\mathbf{s}$, in the constellation vector of size $M = 2^m$. In this paper, the modulation scheme employed is Gray-labelled M-ary quadrature amplitude modulation (QAM) for the transmitted symbols from the nodes a and b. These complex symbols of a particular transmitter node are defined as $\mathbf{s}_i = [s_i^1, s_i^2, \ldots, s_i^{N_{tx}}]^T \in \mathbb{C}^{N_{tx}\times 1}$ where $i \in \{a, b\}$, in which $\mathbb{E}[\|\mathbf{s}_i\|^2] = E_s/N_{tx}$ and $\mathbb{E}[\mathbf{s}_i \mathbf{s}_i^{\mathcal{H}}] = E_s/N_{tx}\mathbf{I}_{N_{tx}}$, where E_s represents the energy of the transmitted symbol. It is worth mentioning that all processing steps applied to the information bits thus far are referred to as BICM .

B. Received Signal Model

At the receiver of node b, after the ADC in each receive antenna branch, the received signal vector, $\boldsymbol{r}_b = [r_1, r_2, \ldots, r_{N_{rx}}]^T \in \mathbb{C}^{N_{rx}\times 1}$, can be given as

$$\boldsymbol{r}_b = \sqrt{\rho_a}\mathbf{H}_{ab}\boldsymbol{s}_a + \sqrt{\widetilde{\rho}_b}\widetilde{\mathbf{H}}_{bb}\boldsymbol{s}_b + \boldsymbol{v}_b, \tag{2}$$

where ρ_a is the average received power from node a, while $\widetilde{\rho}_b$ represents the power from node b's transmitter port to its receiver port after implementing the passive suppression of SI along with the analog/digital SIC techniques, as will discussed later on in Subsection II-D. It is noteworthy that $\widetilde{\mathbf{H}}_{bb}$ represents here the effective SI channel in the digital domain after taking into account the effects of passive SIC methods along with the analog filter, $\mathbf{C}_a$, and digital filter, $\mathbf{C}_d$, which will be discussed later in this paper. In other words, we can define the effective channel of the residual SI in the digital domain, $\widetilde{\mathbf{H}}_{bb}$, as $\widetilde{\mathbf{H}}_{bb} = f(\mathbf{H}_{bb}, \text{Passive SIC}, \mathbf{C}_\text{a}, \mathbf{C}_\text{d})$. Furthermore, it is assumed in this paper that the variances of the noise and the channel state information (CSI) are perfectly known to the receiver.

C. Channel Estimation for FD-MIMO Transceivers

Although the channel estimation is out of the scope of this paper, we briefly outline techniques that can be exploited for this purpose. Achieving high transmission rates in systems operating in FD mode requires precise knowledge of the desired and SI channels [31]. One of the simple and reliable approaches, which can be utilized for this purpose is the pilot-aided technique. Known data symbols, referred to as pilots, can be transmitted during the training mode and algorithms such as least-square (LS) or maximum-likelihood (ML) can utilize the known training pilots to estimate the required channel matrix [12], [27]. During training the system typically operates in HD mode. In the FD-MIMO case, a sequence of known independent pilots, which contains N_p samples, is transmitted from each antenna.

In [12], an efficient protocol is proposed for an FD-MIMO bidirectional transceiver to estimate the SI channel matrices $\mathbf{H}_{ii}$, $\mathbf{H}_{jj}$, and the desired channels $\mathbf{H}_{ij}$, where i and $j \in \{a, b\}$. In this protocol, the time specified for channels estimation is divided into two sub-periods, $\tau_p(1)$ and $\tau_p(2)$, in which the transmitter of node a transmits known pilots during $\tau_p(1)$, while the transmitter of node b remains in the silent mode. Conversely, during the second sub-period, $\tau_p(2)$, node b starts to send its pilot symbols, while the transmitter of node a is silent. The pilot sequence can be defined as $\mathbf{S}_i = [\mathbf{s}_i(1), \mathbf{s}_i(2), \ldots, \mathbf{s}_i(TN_{tx})] \in \mathbb{C}^{N_{tx}\times TN_{tx}}$, where T represents a discrete time interval in which each of $\tau_p(1)$ and $\tau_p(2)$ are sampled. The selection of $\mathbf{S}_i$ should satisfy the condition $\frac{1}{2T}\mathbf{S}_i\mathbf{S}_i^{\mathcal{H}} = \mathbf{I}_{N_{tx}}$, in which the scaling is used for the purpose of satisfying a power constraint over a particular period.

The received training pilots can be expressed as [12]

$$\mathbf{R}_i = \sqrt{\rho_i}\mathbf{H}_{ij}(\mathbf{S}_j + \Omega_j) + \Phi_i + \mathbf{V}_i, \tag{3a}$$

$$\mathbf{R}_j = \sqrt{\rho_j}\mathbf{H}_{jj}(\mathbf{S}_j + \Omega_j) + \Phi_j + \mathbf{V}_j, \tag{3b}$$

where $\boldsymbol{\Omega}$ and Φ are the transmitter noise and the entire distortions induced in the receive chain, respectively. Additionally, $\mathbf{V}$ is the AWGN at the receiver. It is noteworthy that (3a) represents the received training pilots at the far-end terminal after passing through the desired channel, while (3b) is to represent these pilots after passing through the loop-back interference channel causing SI at the same node. Finally, by using LS algorithm the estimate matrices of the desired and the SI channels are respectively expressed as [12]

$$\widehat{\mathbf{H}}_{ij} = \frac{1}{T\sqrt{\rho_i}}\mathbf{R}_i\mathbf{S}_j^{\mathcal{H}}, \tag{4a}$$

$$\widehat{\mathbf{H}}_{jj} = \frac{1}{T\sqrt{\rho_j}}\mathbf{R}_j\mathbf{S}_j^{\mathcal{H}}. \tag{4b}$$

D. Self-Interference Cancellation Stages

As mentioned briefly in the Introduction Section, SI can be combated by utilizing different stages of passive and/or active SIC. This is because SI power may reach or exceed by 100 dB the power of the desired received signal [34], [35]. Thus, our aim is to attenuate the power of SI to be approximately at the level of noise floor, which is approximately -90 dBm [33] as shown in Fig. 2 in [35].

To the best of our knowledge, SI attenuations of about 80 to 120 dB have been reported in the literature [13], [34]–[36] using different stages of SIC. However, in practice, residual SI may still be significant after SIC due to the fact that it is impossible to model perfectly the impairments in the transmit and receive chains, especially the underlying unknown linear and non-linear phenomena. Therefore, we propose IDD in this paper to further reduce the residual SI that remains after applying passive and active SIC approaches.

1) Passive Suppression via Antenna Separation: In order to suppress SI in FD systems, it is initially required to reduce the effects of local power coupling to avoid drowning of the desired incoming signal in loop-interference, which is

significantly stronger. Additionally, it is required to mitigate saturation of ADC circuitry due to its limited dynamic range and quantization resolution [3], [9], [11], [12]. Thus, passive suppression has been proposed at the receiver front-end by using natural-isolation techniques, or sometimes called passive-suppression, via antenna separation to diminish and block the line-of-sight (LoS) path. This is achieved by orienting transmit antenna elements to an opposite direction than those of the receiving antennas, which consequently maximizes LI attenuation by increasing the insertion loss that can be further increased by utilizing orthogonal polarization schemes [3], [9], [13].

2) Analog Domain Cancellation: In this stage, it is required to design the analog cancellation filter, $\mathbf{C}_a \in \mathbb{C}^{N_{tx}\times N_{rx}}$, and combine it with the analog transmitted signal which is perfectly known to the transceiver. This is in order to create an analog replica of SI and apply a subtraction before ADC. In theory, SI can be entirely removed by choosing $\mathbf{C}_a = -\mathbf{H}_{bb}$. However, it is not possible in practice to achieve perfect cancellation [26], [13]. In other words, a precise implementation of this filter is not easily within reach in the analog domain, as it is required to be abruptly adaptive to any change in the wireless channels. Moreover, to cancel the strongest propagation path, only a filter matrix with dimensions (N_{tx}, N_{rx}) is required for the implementation of the corresponding analog filters [26], which is not adequate to cancel SI totally in our case in this paper for MIMO system. This stage is still significant as it produces, along with antennas separation, an SI mitigation which is required firstly to avoid the saturation at the received port front end and secondly to overcome the quantization noise in the ADC caused by large magnitude of SI with compare to the signal of interest [13].

3) Digital Domain Cancellation: Digital cancellation is utilized to suppress the residual SI passed through ADC to the digital domain. The digital filter, $\mathbf{C}_d \in \mathbb{C}^{N_{tx}\times N_{rx}}$, is designed to perform this task by taking into account the interfering channel, the analog filter, and the ADC processing. This filter can implemented to remove the residual SI as $\mathbf{C}_d = -\mathbf{A}(\mathbf{H}_{bb} + \mathbf{C}_a)$, where $\mathbf{A}$ represents a real diagonal matrix resulting from the ADC process [26]. However, this stage of SI cancellation does not have the ability to remove all parts of residual SI coming from the analog domain, in addition to the clipping-plus-quantization noise terms caused by ADC proceeding. Therefore, an additional signal processing stage is required to reduce the effect of residual SI. In this paper, we propose IDD for coded FD-MIMO for this purpose as will be discussed in the next subsection.

E. Iterative Detection and Decoding

The key novelty and main challenge addressed in our paper is the derivation of the tight upper bound performance for this type of self-interference characteristics, which is distinct from the propagation characteristics of the distant user. To the best of our knowledge, the derived tight upper bound for this scenario is not known in the available literature. Furthermore, its derivation is not simple and deserves to be disseminated since it is a useful tool for full-duplex modem designers to predict the performance of these systems.

Fig. 2 shows the block diagram of IDD with soft-PIC for node b. It is worth noting that an identical type of receiver is used at node a. The detector comprises an adaptive MMSE filter, which is discussed later in this subsection, and an LLR demapper. A soft decoder performs soft channel decoding for the extrinsic information from the LLR demapper. During the first iteration the detector and decoder need to be initialized since the *a priori* and *a posteriori* information has not been obtained yet. Therefore, at that instant, the soft estimations of the SI symbols are not available, and thus, they cannot be subtracted from the received signal [8]. This means that the received MIMO signal, $\boldsymbol{r}_b$, for this iteration, could be considered as the input for the MMSE filter, rather than the reduced interference signal, $\widetilde{\boldsymbol{r}}_b$, obtained in the next iterations. For the same reasons mentioned above and due to the lack of the required parameters, the MMSE filter is not ready to achieve interference cancellation during the first iteration; therefore, it performs classical MMSE filtering and passes its outputs, $\boldsymbol{y}$, to the LLR demapper [8]. The latter first transforms the filtered symbols to LLR symbols and performs a soft demapping of these symbols to bit-wise LLRs. This operation can be expressed mathematically as

$$\mathcal{L}'[c_n] = \ln \frac{Pr\left\{c_n = 1 \mid \boldsymbol{y}_b\right\}}{Pr\left\{c_n = 0 \mid \boldsymbol{y}_b\right\}} \tag{5a}$$

$$= \ln \frac{Pr\left\{c_n = 1 \mid \widetilde{\boldsymbol{r}}_b, \boldsymbol{w}_b\right\}}{Pr\left\{c_n = 0 \mid \widetilde{\boldsymbol{r}}_b, \boldsymbol{w}_b\right\}}, \tag{5b}$$

where (5b) represents the LLR of each coded bit with respect to the MMSE filter output $\boldsymbol{y}_b = \boldsymbol{w}_b^{\mathcal{H}}\, \widetilde{\boldsymbol{r}}_b$, with $\widetilde{\boldsymbol{r}}_b$ and $\boldsymbol{w}_b$ being its input and filter weights, respectively. The soft LLR bits with a size of $(N_{tx} \times m N_{symbol})$, where N_{symbol} denotes the number of symbols per frame, are converted from a parallel to a serial stream, deinterleaved and processed by the soft decoder. This *a priori* information provided to the decoder is processed by exploiting the linear to logarithmic approximation of the maximum *a posteriori* algorithm (linear-log-MAP), as utilized in this paper. This operation can be written in mathematical form as

$$\mathcal{L}''[c_n] = \ln \frac{Pr\left\{c_n = 1 \mid \mathcal{L}'[c_n]\right\}}{Pr\left\{c_n = 0 \mid \mathcal{L}'[c_n]\right\}}. \tag{6}$$

At this stage, a soft estimation for all symbols from the decoder output can be obtained as [10]

$$\widehat{\boldsymbol{s}}_a \triangleq \mathbb{E}(\mathbf{s}_a) = \sum_{\alpha_t \in \mathcal{S}} \alpha_t Pr\left\{s_a = \alpha_t\right\}, \tag{7}$$

in which for $M = 4$, i.e. 4-QAM, the soft remapping of the symbols can be expressed as

$$\widehat{\boldsymbol{s}}_a = \frac{1}{\sqrt{2}}\left[\tanh\left(\frac{\tilde{\mathcal{L}}''[c_{n,1}]}{2}\right) + j\tanh\left(\frac{\tilde{\mathcal{L}}''[c_{n,2}]}{2}\right)\right], \tag{8}$$

where $\tilde{\mathcal{L}}''$ represents the LLR symbols after re-interleaving, i.e. $\tilde{\mathcal{L}}''[.] = \Pi\Big(\mathcal{L}''[.]\Big)$. Additionally, $c_{n,1}$ and $c_{n,2}$ denote the successive odd and even re-interleaved coded bits, respectively.

On the other hand, the soft remapping of 16-QAM can be expressed as [37], [38]

$$\widehat{s}_a = \frac{1}{\sqrt{2.5}}\Big[\big(1-2Pr\{c_{n,1}=0\}\big)\big(1+2Pr\{c_{n,1}=0\}\big) + j\big(1-2Pr\{c_{n,3}=0\}\big)\big(1+2Pr\{c_{n,4}=0\}\big)\Big], \quad (9)$$

where $Pr\{c_{n,m}=0\}$ is the probability of the m^{th} coded bit belonging to a 16-QAM symbol of the n^{th} antenna. Moreover, $Pr\{c_{n,m}=0\}$ can be determined as [37]

$$Pr\{c_{n,m}=0\} = \frac{1}{2}\left(1+\frac{1}{2}\tanh\left(\tilde{\mathcal{L}}'[c_{n,m}]\right)\right). \quad (10)$$

The estimated symbols, $\widehat{s}_{a,n}$, are used along with CSI to create a soft replica of the received signal along with its interference. It is assumed that perfect CSI is available, since channel estimation is beyond the scope of this paper. The estimated symbols are then reshaped into a parallel vector in order to make each symbol correspond to its original antenna element, and subsequently the estimated symbols associated with the transmitted n^{th} antenna element are forced to zero, as this is required to cancel its induced interference [8], i.e.

$$\underline{\widehat{\boldsymbol{s}}}_{a,n} \triangleq \left\{\widehat{s}_{a,1},\ldots,\widehat{s}_{a,n-1},0,\widehat{s}_{a,n+1},\ldots,\widehat{s}_{a,N_{tx}}\right\}. \quad (11)$$

In other words, $\underline{\widehat{\boldsymbol{s}}}_{a,n}$ represents a vector of the soft estimated symbols of the transmitted antennas after nullifying the symbol of the n^{th} antenna, which is the subject of interest and it is at this instant subject to the soft-PIC processing [8]. The soft symbols in (11) are used to create an estimate of the transmitted signal with the interference after combining them with the channel, i.e. $\widehat{\boldsymbol{r}}_b = \mathbf{H}_{ab}\underline{\widehat{\boldsymbol{s}}}_{a,n}$. Subsequently, soft-PIC is performed by subtracting $\widehat{\boldsymbol{r}}_b$ from the received signal as

$$\widetilde{\boldsymbol{r}}_b = \boldsymbol{r}_b - \widehat{\boldsymbol{r}}_b = \mathbf{H}_{ab}\boldsymbol{s}_{a,n} + \widetilde{\mathbf{H}}_{bb}\boldsymbol{s}_{b,n} + \boldsymbol{v}_b - \mathbf{H}_{ab}\underline{\widehat{\boldsymbol{s}}}_{a,n} \quad (12a)$$

$$= \mathbf{H}_{ab}\widetilde{\boldsymbol{s}}_{a,n} + \widetilde{\mathbf{H}}_{bb}\boldsymbol{s}_{b,n} + \boldsymbol{v}_b. \quad (12b)$$

where $\widetilde{\boldsymbol{s}}_{a,n}$ is defined as

$$\widetilde{\boldsymbol{s}}_{a,n} \triangleq \left\{\widetilde{e}_{a,1},\ldots,\widetilde{e}_{a,n-1},s_{a,n},\widetilde{e}_{a,n+1},\ldots,\widetilde{e}_{a,N_{tx}}\right\}, \quad (13)$$

and $\widetilde{\boldsymbol{e}}_{a,n} = \boldsymbol{s}_{a,n} - \underline{\widehat{\boldsymbol{s}}}_{a,n}$ represents the error which arises due to the imperfect cancellation of interference on node b. This error might theoretically approach to zero when the number of iterations approaches infinity.

At the end of the first iteration, and for the iterations beyond, the information required is provided to the adaptive MMSE filter in order to perform further cancellation of the residual SI. This filter is designed to achieve a minimization of the mean squared error (MSE) between its output, $s^a_{n,l}$, with respect to the l^{th} symbol of the n^{th} transmitted antenna of node a, i.e.

$$\mathbf{J}_{n,l} = \mathbb{E}\left\|\boldsymbol{w}^{\mathcal{H}}_{n,l}\widetilde{\boldsymbol{r}}^b_{n,l} - \boldsymbol{s}^a_{n,l}\right\|^2. \quad (14)$$

The coefficient w that minimizes $\mathbf{J}_{n,l}$ can be obtained by solving $\frac{\partial \mathbf{J}_{n,l}}{\partial \boldsymbol{w}^b_{n,l}} = \mathbf{0}$, and was derived in [39] and also given in [8], [19], and [24] for different types of interference. Following the derivation in [31], the solution in the case of SI can be given as

$$\boldsymbol{w}^b_{n,l} = [\Lambda_s(l)\mathbf{H}_{ab}(l)^{\mathcal{H}}\mathbf{H}_{ab}(l) + \Lambda_i(l)\widetilde{\mathbf{H}}_{bb}(l)^{\mathcal{H}}\widetilde{\mathbf{H}}_{bb}(l) + \sigma^2_{v_b}\mathbf{I}_{Nrx}]^{-1}\boldsymbol{h}_{ab}(n,l)^{\mathcal{T}}, \quad (15)$$

where $\boldsymbol{h}_{ab}(n,l)$ represents the n^{th} column of $\mathbf{H}_{ab}(l)$. Furthermore, $\Lambda_i(l)$ denotes the covariance matrix of the SI symbols, $\boldsymbol{s}_{b,n}$, which can be expressed as $\Lambda_i(l) = \mathbb{E}[\boldsymbol{s}_{b,n}\boldsymbol{s}^{\mathcal{H}}_{b,n}]$, while $\Lambda_s(l) = \mathbb{E}[\widetilde{\boldsymbol{s}}_{a,n}\widetilde{\boldsymbol{s}}^{\mathcal{H}}_{a,n}]$ represents the covariance matrix of $\widetilde{\boldsymbol{s}}_{a,n}$ defined as

$$\Lambda_s(l) \triangleq \mathrm{diag}\Big[\mathbb{E}[\|\widetilde{e}_{a,1}\|^2],\ldots,\mathbb{E}[\|\widetilde{e}_{a,n-1}\|^2],\sigma^2_{s_a}, \mathbb{E}[\|\widetilde{e}_{a,n+1}\|^2],\ldots,\mathbb{E}[\|\widetilde{e}_{a,N_{tx}}\|^2]\Big]. \quad (16)$$

Using the orthogonality principle, i.e. $\mathbb{E}[(\boldsymbol{w}^{\mathcal{H}}_{n,l}\widetilde{\boldsymbol{r}}^b_{n,l} - \boldsymbol{s}^a_{n,l})(\widetilde{\boldsymbol{r}}^b_{n,l})^{\mathcal{H}}] = 0$, between the input of the MMSE filter, $\widetilde{\boldsymbol{r}}^b_{n,l}$ and the MSE, it can be shown that the MMSE, $\mathbf{J}^{\min}_{n,l}$, achieved by using the coefficients in (15) is given as

$$\mathbf{J}^{\min}_{n,l} = \Lambda_s(l) - \mathbf{R}_{rs}\mathbf{R}_r^{-1}\mathbf{R}_{sr}, \quad (17)$$

where $\mathbf{R}_{rs} = \mathbb{E}[\widetilde{\boldsymbol{r}}^b_{n,l}(\widetilde{\boldsymbol{s}}_{a,n})^{\mathcal{H}}]$, $\mathbf{R}_r = \mathbb{E}[\widetilde{\boldsymbol{r}}^b_{n,l}(\widetilde{\boldsymbol{r}}^b_{n,l})^{\mathcal{H}}]$, and $\mathbf{R}_{sr} = \mathbb{E}[\widetilde{\boldsymbol{s}}_{a,n}(\widetilde{\boldsymbol{r}}^b_{n,l})^{\mathcal{H}}]$.

Since the MAP soft decoder is used in this paper, $\mathbb{E}[\|\widetilde{e}_{a,n}\|^2]$ can be computed as

$$\mathbb{E}[\|\widetilde{e}_{a,n}\|^2] = \sum_{\alpha_t\in\mathcal{S}}\|\boldsymbol{s}_{a,n} - \underline{\widehat{\boldsymbol{s}}}_{a,n}\|^2 Pr\left\{s_{a,n}=\alpha_t\right\}, \quad (18)$$

where $\sigma^2_{s_a} = \mathbb{E}[\|\boldsymbol{s_{a,n}}\|^2] = 1$ represents the variance of the symbols $\boldsymbol{s_{a,n}}$ as defined above in Section II-A.

Once the computation of the adaptive MMSE filter coefficients, $\boldsymbol{w}^b_{n,l}$, is complete, they are used along with output of the soft-PIC, $\widetilde{\boldsymbol{r}}^b_{n,l}$, to obtain the input of the LLR demapper as

$$\boldsymbol{y}^b_{n,l} = (\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\,\widetilde{\boldsymbol{r}}^b_{n,l} \quad (19a)$$

$$= \beta_{n,l}\,s^a_{n,l} + \zeta_{n,l}, \quad (19b)$$

where $\beta_{n,l} = (\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\boldsymbol{h}_{ab}(n,l)$, while $\zeta_{n,l}$ represents the residual-SI-plus-noise term, which is defined as

$$\zeta_{n,l} = (\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\widetilde{\boldsymbol{h}}_{bb}(n,l)\,s^b_{n,l} + (\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\boldsymbol{v}_b, \quad (20)$$

where $\widetilde{\boldsymbol{h}}_{bb}(n,l)$ represents the n^{th} column of $\widetilde{\mathbf{H}}_{bb}(l)$. Moreover, $\zeta_{n,l}$ in (20) exhibits a Gaussian distribution with zero-mean and variance derived in [39] as

$$\sigma^2_{\zeta_{n,l}} = \frac{E_s}{N_{tx}}(\beta_{n,l} - \beta^2_{n,l}). \quad (21)$$

Furthermore, the residual interference-to-noise ratio (INR), $\Omega_{n,l}$, can be obtained from (20) as

$$\Omega_{n,l} = \frac{\left\|(\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\widetilde{\boldsymbol{h}}_{bb}(n,l)\,s^b_{n,l}\right\|^2}{\left\|(\boldsymbol{w}^b_{n,l})^{\mathcal{H}}\boldsymbol{v}_b\right\|^2}. \quad (22)$$

At the output of the MMSE filter, the signal-to-interference-plus-noise ratio (SINR) for the l^{th} spatial symbol can be defined similar to [40] as

$$\gamma_l = \frac{\beta_l}{(1-\beta_l)}, \quad (23)$$

where the distribution of γ_l is derived in Appendix by employing the moment generation function (MGF).

All of these signal processing steps discussed so far for the first iteration are repeated for a predefined number of iterations so that the system converges to an acceptable bit-error rate (BER) with respect to SINR [8]. After the last iteration is

completed, a hard decision is applied to the soft decoder output in order to obtain the transmitted bits as

$$\widehat{a}_q^a = \text{sign}\Big(\mathcal{L}''\big[a_q^a\big]\Big), \tag{24}$$

where

$$\mathcal{L}''[a_q^a] = \ln \frac{Pr\left\{a_q^a = 1 \mid \mathcal{L}'[c_p^a]\right\}}{Pr\left\{a_q^a = 0 \mid \mathcal{L}'[c_p^a]\right\}}. \tag{25}$$

III. Tight Upper Bound Performance Analysis

In this section, we derive a tight upper bound on the performance of the proposed coded FD-MIMO-BICM system in the presence of residual SI that remains after applying the passive and active SIC methods mentioned previously in this paper. This bound is derived by assuming Rayleigh fading distributions for desired and SI channels. Moreover, the derived bound is valid under the assumption of Rician distribution of the SI channel with small κ-factor, due to imperfect suppression of the LoS path. In general, the BICM bound of the BER, $\bar{P}_b$, for a rate-1/n convolutional code is given by [41, p. 515] and [42] as

$$\bar{P}_b \leq \sum_{d=d_{free}} B_d P_d\left(\mathbf{s}, \mathbf{y}\right), \tag{26}$$

where $P_d\left(\mathbf{s}, \mathbf{y}\right)$ denotes the pair-wise error probability (PEP) between the transmitted and estimated sequences $\mathbf{s} = \left\{s_1, s_2, \ldots, s_{N_{tx}}\right\}$ and $\boldsymbol{y} = \left\{y_1, y_2, \ldots, y_{N_{tx}}\right\}$, respectively, when the Hamming distance between them is equal to d. Additionally, d_{free} is the convolutional code's free distance, B_d represents the total input weight of error events at Hamming distance d. These coefficients, which are presented in [42] for different R_c in tables, can be obtained by applying the derivation to the transfer function of the code, $T\left(B, D\right)$, as

$$\left.\frac{\partial T\left(B, D\right)}{\partial B}\right|_{B=1} = \sum_{d=d_{free}}^{\infty} B_d D^d, \tag{27}$$

where D and B represent the Hamming distances of the coded and input sequences, respectively.

According to (19b), we can rewrite (5) for the i^{th} $(i \in \left\{1, 2, \ldots, \log_2 M\right\})$ coded bit, in the l^{th} symbol belonging to the n^{th} transmit antenna element of $s_{n,l}^a$ with respect to $y_{n,l}^a$ as

$$\begin{aligned}
\mathcal{L}'(n,l,i) &= \ln \frac{\sum\limits_{\tilde{s} \in \mathcal{S}_{n,l,i}^1} Pr(y_{n,l}^b | \tilde{s}, \beta_{n,l}, \zeta_{n,l})}{\sum\limits_{\tilde{s} \in \mathcal{S}_{n,l,i}^0} Pr(y_{n,l}^b | \tilde{s}, \beta_{n,l}, \zeta_{n,l})}, \\
&= \ln \left(\frac{\sum\limits_{\tilde{s} \in \mathcal{S}_{n,l,i}^1} e^{-\frac{|\beta_{n,l}(s_{n,l}^a - \tilde{s}) + \zeta_{n,l}|^2}{\sigma_{\zeta_{n,l}}^2}}}{\sum\limits_{\tilde{s} \in \mathcal{S}_{n,l,i}^0} e^{-\frac{|\beta_{n,l}(s_{n,l}^a - \tilde{s}) + \zeta_{n,l}|^2}{\sigma_{\zeta_{n,l}}^2}}} \right), \\
&= \ln(\Xi_{n,l}),
\end{aligned} \tag{28}$$

where $\mathcal{S}_{n,l,i}^1$ and $\mathcal{S}_{n,l,i}^0$ represent the signal subsets in the constellation when the i^{th} bit in the l^{th} symbol and for the n^{th} transmit antenna is equal to 1 and 0, respectively. We assume that a message of all zeros is transmitted, then the PEP is expressed as

$$P_d\left(\mathbf{s}, \mathbf{y}\right) = Pr\Big(\sum_{k=1}^{d} \mathcal{L}_k'\Big), \tag{29}$$

where $\mathcal{L}_k'$ is the input to the soft decoder after applying parallel-to-serial conversion and deinterleaving to $\mathcal{L}'(n,l,i)$. Since it is not straightforward to evaluate the distribution of $\mathcal{L}_k'$ in an exact expression, we can exploit the MGF approach, denoted as $\mathcal{M}$, to evaluate the probability in (29) as [41]

$$P_d\left(\mathbf{s}, \mathbf{y}\right) = \int_{\delta - j\infty}^{\delta + j\infty} \mathcal{M}_{\sum_1^d \mathcal{L}_k'}(t) \frac{dt}{t} \tag{30a}$$

$$= \int_{\delta - j\infty}^{\delta + j\infty} \left[\mathcal{M}_{\mathcal{L}'}(t)\right]^d \frac{dt}{t}, \tag{30b}$$

where (30b) is obtained from the assumption of random interleaving. Moreover, the subscript k is omitted on $\mathcal{L}'$ in (30b) since the statistics of LLR are calculated at a single instant in time. Additionally, δ is a constant that can be obtained by minimizing the value of $\mathcal{M}_{\mathcal{L}'}(t)$. The procedure to obtain the optimum value of δ is described later on in this paper. Furthermore, $\mathcal{M}_{\mathcal{L}'}(t)$ can be evaluated similar to [43] and [40] as

$$\begin{aligned}
\mathcal{M}_{\mathcal{L}'}(t) &= \mathbb{E}_{s_{n,l}^a, \beta_{n,l}, \zeta_{n,l}}\Big[\exp(t\mathcal{L}')\Big] \\
&= \frac{1}{M\,m} \sum_{t=1}^{m} \sum_{s_{n,l}^a \in \mathcal{S}} \mathcal{I}_{s_{n,l}^a}(t),
\end{aligned} \tag{31}$$

where $m = \log_2(M)$ and $\mathcal{S}$ represents a set of all possible complex symbols of the M-QAM constellation. Additionally, $\mathcal{I}_{s_{n,l}^a}(t)$ is expressed as

$$\mathcal{I}_{s_{n,l}^a}(t) = \mathbb{E}_{\beta_{n,l}, \zeta_{n,l}}\Big[e^{t \ln(\Xi_{n,l})}\Big], \tag{32}$$

which can be simplified as

$$\mathcal{I}_{s_{n,l}^a}(t) = \mathbb{E}_{\beta_{n,l}, \zeta_{n,l}}\Big[\left(\Xi_{n,l}\right)^t\Big]. \tag{33}$$

A closer look to the ratio in (33) reveals that it is dominated by a single term representing the minimum distance in the numerator and denominator, conditioned to high SINR [43]. According to the assumption of transmitting an all-zero message and by utilizing the theorem of dominated convergence [44], (33) can be approximated as

$$\mathcal{I}_{s_{n,l}^a}(t) \simeq \mathbb{E}_{\beta_{n,l}, \zeta_{n,l}}\left[e^{\frac{t|\zeta_{n,l}|^2 - t|\beta_{n,l}(s_{n,l}^a - \tilde{s}) + \zeta_{n,l}|^2}{\sigma_{\zeta_{n,l}}^2}}\right], \tag{34}$$

where $\tilde{s} \in \mathcal{S}_{n,l,i}^1$ represents the closest symbol to $s_{n,l}^a$ in the M-QAM constellation.

TABLE I
RATE-1/2 RSC CHANNEL ENCODERS PARAMETERS USED IN (36)

	$\Delta_{M,k}$	$\Upsilon_{M,k}$
4-QAM	{1}	{2.0}
16-QAM	{0.75, 0.25}	{0.4, 1.6}

At this stage, we resume the derivation by simplifying (34) and evaluate its average over $\zeta_{n,l}$ to obtain

$$\mathcal{I}_{s^a_{n,l}}(t) \simeq \mathbb{E}_{\beta_{n,l}}\left[\exp\left(-\frac{t(1-t)N_{tx}}{E_s}\gamma_l\left|s^a_{n,l}-\tilde{s}\right|^2\right)\right] \simeq \mathcal{M}_{N_{rx},N_{tx}}\left[\exp\left(\frac{t(1-t)N_{tx}}{E_s}\left|s^a_{n,l}-\tilde{s}\right|^2\right)\right], \quad (35)$$

where $\mathcal{M}_{N_{rx},N_{tx}}$ represents the MGF of the SINR, γ_l, which is derived in Appendix. Now, by substituting (35) in (31) and using the simplifications described in [43], we can rewrite (31) as

$$\mathcal{M}_{\mathcal{L}'}(t) = \sum_k \Delta_{M,k}\mathcal{M}_{N_{tx},N_{rx}}\left(\frac{t(1-t)N_{tx}}{E_s}\Upsilon_{M,k}\right), \quad (36)$$

where $\Upsilon_{M,k}$ and $\Delta_{M,k}$ are the squared Euclidean distances and the frequency of the occurrence for each distance presented in Table I [43], respectively.

Now, we can substitute (36) in (30b) and apply the Gauss-Chebyshev quadrature (GCQ) rule in order to obtain the PEP of the proposed system as

$$P_d(\mathbf{s},\mathbf{y}) = E_\varphi + \frac{1}{\varphi}\sum_{k=1}^{\varphi/2}\left(\Re\left[\mathcal{M}_{\mathcal{L}'}(\epsilon)^d\right] + \tau_k\Im\left[\mathcal{M}_{\mathcal{L}'}(\epsilon)^d\right]\right), \quad (37)$$

where $\epsilon = \delta + j\delta\tau_k$, $\tau_k = \tan\left(\frac{(2k-1)\pi}{2\varphi}\right)$, φ represents the number of nodes utilized in applying the GCQ to (36), and E_φ is an integration constant approaching zero when φ approaches infinity.

As previously mentioned, δ is a constant chosen to minimize the value of $\mathcal{M}_{\mathcal{L}'}(t)$. Thus, from (35) and the definition of MGF in (49) in Appendix, and by setting the result of the derivative with respect to the variable $-t(1-t)$ to zero, i.e. $\frac{d[-t(1-t)]}{dt} = 0$, it can be shown that the minimum value of $\mathcal{M}_{\mathcal{L}'}(t)$ is obtained at $t = 0.5$. Additionally, since $\mathcal{M}_{N_{tx},N_{rx}}$ is a monotonic increasing function, $\mathcal{M}_{\mathcal{L}'}(t)$ exhibits a minimum value at $t = 0.5$. It is noteworthy that $\mathcal{M}_{\mathcal{L}'}(\delta + j\delta\tau_k)$ is real-valued when $\Re[\epsilon] = \delta = 0.5$. Therefore, the approximate formula of the PEP in (37) can be written as

$$P_d(\mathbf{s},\mathbf{y}) \simeq \frac{1}{\varphi}\sum_{k=1}^{\varphi/2}\left[\sum_i \Delta_{M,i}\mathcal{M}_{N_{tx},N_{rx}}\left(\frac{(1+\tau_k^2)N_{tx}}{4E_s}\Upsilon_{M,i}\right)\right]^d. \quad (38)$$

Fig. 3 reveals a close agreement of the simulated BER performance as a function of SNR for the 2nd iteration with the tight upper bounds performance for the three cases of 1/2-rate RSC channel encoders, whose parameters are illustrated in Table II, and for the case of $N^a_{tx} = 2, N^b_{rx} = 4, N^b_{tx} = 2$ in the presence of a residual interference-to-noise ratio (INR) of 10 dB. The modulation scheme used in this figure is 4-QAM. Additionally, it is worth mentioning that in order to achieve an acceptable precision, the number of nodes, φ, chosen for the GCQ rule in (38) is 25.

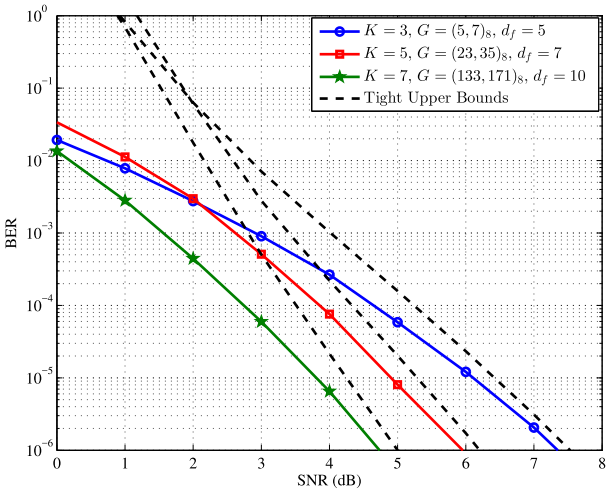

Fig. 3. BER vs. SNR performance after the 2nd iteration along with the tight upper bounds for different 1/2-rate convolutional codes of (2,4,2)-FD-MIMO-BICM-IDD utilizing 4-QAM in the presence of 10 dB residual INR.

A Rayleigh fading distribution is assumed for all the channels in this paper except the SI channel. For the SI channel, we assume a Rician distribution with a small κ-factor. We examine our tight upper bound performance for the case when the SIC approaches are unable to perfectly remove the LoS path [33]. Fig. 4 demonstrates the performance of the proposed system for this scenario using different combinations of $(2, N_{rx}, 2)$, i.e. $N_{rx} = 4, 5, 6$, for rate-1/2 RSC codes with constraint length $K = 5$ as shown in Table II, and for the case of 4-QAM modulation scheme. Additionally, the case of Rayleigh distribution of the SI channel is included in this figure too, when $\kappa = 0$ by assuming perfect LoS suppression. This figure reveals that our tight upper bound performance is valid for Rician distributions of SI channel when the κ-factor is below 1 dB after 5 IDD iterations.

IV. EXIT CHART ANALYSIS

EXIT chart is a semi-analytical technique that can be exploited to visualize the convergence behavior of iterative detection and decoding systems. It was proposed in [45] as a powerful tool to analyze and characterize the flow of the mutual information (MI), which is exchanged between the constituent detector and decoder in an iterative system. In this paper, the MI terms at the input and output of the soft decoder are denoted as I^i_D and I^o_D, respectively, while I^i_E and I^o_E represent the MI terms at the input and output of the equalizer, respectively. These MI terms are evaluated as [46]

$$I_D = \frac{1}{2}\sum_{c\in\{\mp 1\}}\int_{-\infty}^{\infty} p_D(\xi|c)\log_2\frac{2p_D(\xi|c)}{p_D(\xi|+1)+p_D(\xi|-1)}d\xi, \quad (39)$$

TABLE II
1/2-RATE RSC CHANNEL ENCODERS PARAMETERS USED IN FIG. 3

K	G	d_f	$\{B_d(d_{free}), B_d(d_{free}+1), ..., B_d(d_{free}+20)\}$
3	$(5,7)_8$	5	{1, 4, 12, 32, 80, 192, 448, 1024, 2304, 5120, 11264, 14576, 53248, 114688, 245760,524288, 1114112,2359296, 4980736, 10485760}
5	$(23,35)_8$	7	{4, 12, 20, 72, 255, 500, 1324, 3680, 8967, 22270, 57403, 142234, 348830, 867106, 2134239,5205290, 12724352, 31022962, 75250693, 182320864}
7	$(133,171)_8$	10	{36, 0, 211, 0, 1404, 0, 11633, 0, 77433, 0, 502690, 0, 3322763, 0, 21292910, 0,134365911, 0, 84342587, 0}

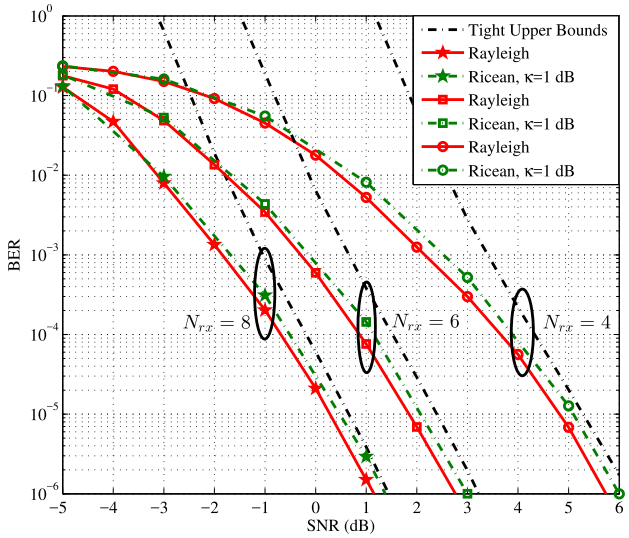

Fig. 4. BER vs. SNR performance of FD-MIMO-IDD after 5 iterations, 4-QAM with 1/2-rate convolutional BICM are used for different combinations of $(2, N_{rx}, 2)$, Rayleigh and small κ-factor Rician distributions are assumed for residual SI channel, and for a residual INR of 10 dB.

and

$$I_E = \frac{1}{2}\sum_{c\in\{\mp 1\}}\int_{-\infty}^{\infty} p_E(\xi|c)\log_2\frac{2p_E(\xi|c)}{p_E(\xi|+1)+p_E(\xi|-1)}d\xi, \tag{40}$$

where $p_D(\xi|c)$ and $p_E(\xi|c)$ are the conditional PDFs of the soft decoder and equalizer, respectively, associated with their *a priori* LLR and the information bits c.

Similar to [45] and [47], we model the *a priori* LLR's of the soft decoder, D, and the equalizer, E as independent Gaussian random variables, n_D and n_E with zero mean and variance σ_D^2 and σ_E^2, respectively. Thus, their *a priori* LLRs can be expressed after joining the information bits, c, as $L_D = \mu_D c + n_D$ and $L_E = \mu_E c + n_E$, where $\mu_D = \sigma_D^2/2$ and $\mu_E = \sigma_E^2/2$. By assuming a sufficiently long interleaver, the general expression for these conditioned PDFs can be given as

$$p_D(\xi|c) = \frac{1}{\sqrt{2\pi}\sigma_D}e^{-\frac{\left(\xi-\frac{\sigma_D^2}{2}c\right)^2}{2\sigma_D^2}}, \tag{41}$$

$$p_E(\xi|c) = \frac{1}{\sqrt{2\pi}\sigma_E}e^{-\frac{\left(\xi-\frac{\sigma_E^2}{2}c\right)^2}{2\sigma_E^2}}. \tag{42}$$

It is worth mentioning that the information bits are assumed to be equiprobable, i.e. $Pr\{c=+1\} = Pr\{c=-1\} = 0.5$.

Now, by substituting (41) in (39) and (42) in (40), I_D and I_E can be re-written as

$$I_D(\sigma_D) = 1 - \frac{1}{\sqrt{2\pi}\sigma_D}\int_{\infty}^{\infty} e^{-\frac{\left(\xi-\frac{\sigma_D^2}{2}\right)^2}{2\sigma_D^2}}\log_2\left(1+e^{-\xi}\right)d\xi, \tag{43}$$

and

$$I_E(\sigma_E) = 1 - \frac{1}{\sqrt{2\pi}\sigma_E}\int_{\infty}^{\infty} e^{-\frac{\left(\xi-\frac{\sigma_E^2}{2}\right)^2}{2\sigma_E^2}}\log_2\left(1+e^{-\xi}\right)d\xi. \tag{44}$$

Alternatively, I_D and I_E can be evaluated by averaging the LLRs as [41, p. 555]

$$\begin{aligned} I_D &= 1 - \mathbb{E}_{c=+1}\Big[\log_2(1+e^{-L_D})\Big], \\ &\simeq 1 - \frac{1}{N}\sum_{n=1}^{N}\log_2(1+e^{-L_D}), \end{aligned} \tag{45}$$

$$\begin{aligned} I_E &= 1 - \mathbb{E}_{c=+1}\Big[\log_2(1+e^{-L_E})\Big], \\ &\simeq 1 - \frac{1}{N}\sum_{n=1}^{N}\log_2(1+e^{-L_E}), \end{aligned} \tag{46}$$

where N here represents the total number of LLRs.

Fig. 5 shows the EXIT chart for the proposed IDD system of $(2, N_{rx}, 2)$-FD-MIMO-BICM for different N_{rx}, i.e. $N_{rx} = 2, 4, 6$, at SNR=-5 dB and residual INR=10 dB. The rate-1/2 RSC with constraint length $K = 5$, which is shown in Table II, is considered and for the case of 4-QAM modulation scheme. Moreover, a Rician distribution with low κ-factor of 1 dB is used to model the SI channel, while the forward channel has a Rayleigh fading distribution. Since the MI at the output of the equalizer I_E^o becomes the MI to the input of the decoder I_D^i and the MI at output of the decoder, I_D^o becomes the MI to the input of the equalizer I_E^i, therefore, the EXIT chart is drawn with two axes, which are $(I_E^o = I_D^i, I_E^i = I_D^o)$. A closer look to Fig. 5 reveals that higher number of receiving antennas, N_{rx}, leads to faster convergence between the soft decoder and the proposed equalizer and additionally reduces the number of iterations required to obtain this convergence. In this figure, the trajectory shows that the number of iterations required for the proposed system for $N_{rx} = 2$ is about 4, while two

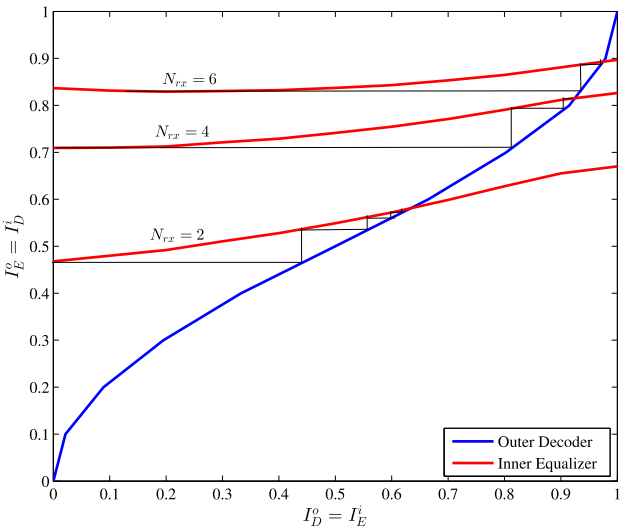

Fig. 5. EXIT chart for a $(2, N_{rx}, 2)$-FD-MIMO system at SNR $= -5$ dB and residual INR of 10 dB. (The black solid lines represent the trajectory).

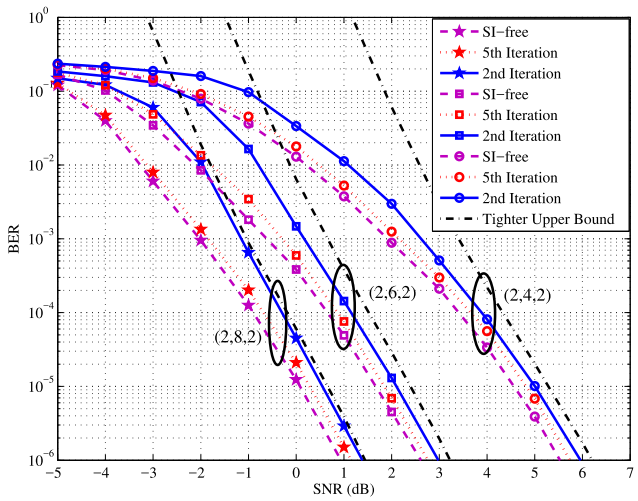

Fig. 6. BER vs. SNR performance of FD-MIMO-IDD utilizing 4-QAM with 1/2-rate convolutional BICM for the 2nd and 5th iterations for different combinations of $\left(N_{tx}^a, N_{rx}^b, N_{tx}^b\right)$ and residual INR=10 dB compared to their corresponding SI-free cases.

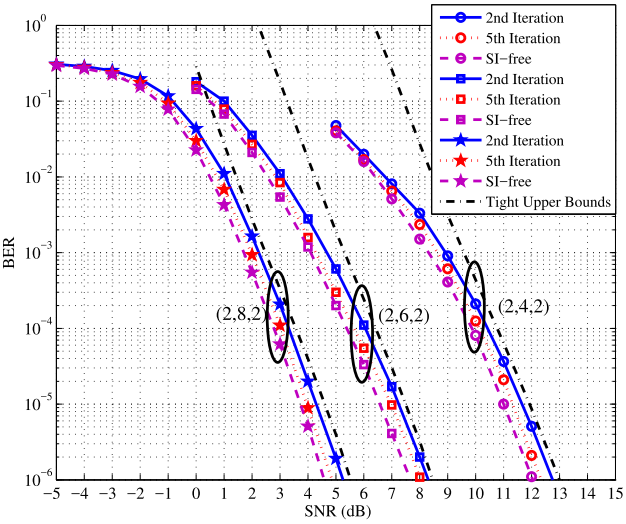

Fig. 7. BER vs. SNR performance of FD-MIMO utilizing 16-QAM with 1/2-rate convolutional BICM for the 2nd and 5th iterations and for different combinations of $\left(N_{tx}^a, N_{rx}^b, N_{tx}^b\right)$ and residual INR=10 dB compared to their corresponding SI-free cases.

iterations are adequate when $N_{rx} = 6$. This demonstrate that the number of the received antennas is inversely proportional to the number of iterations required to achieve converge.

V. Simulation Results and Discussion

In this section, the performance of an FD-MIMO based bi-directional transceiver of a two-node system, a and b, is considered. Performance is impaired by self-interference and mitigated using IDD, which firstly performs soft-PIC to mitigate the SI and then adaptive MMSE filtering to remove the residual interference. Moreover, in all the presented simulations, BICM is utilized with 1/2-rate RSC codes with the constraint length of $K = 5$, which its parameters are illustrated in the second raw of Table II. The overall system performance is evaluated in the presence of AWGN over frequency non-selective Rayleigh fading channels. Furthermore, the modulation scheme used in the simulations is 4-QAM and 16-QAM in conjunction with spatial multiplexing MIMO. We assume 40 dB of passive suppression by antenna separation followed by two stages of analog and digital filtering achieving 70 dB of additional SI attenuation [36].

Figs. 6 and 7 show the performance of the proposed system by utilizing two different M-QAM modulation schemes which are 4-QAM and 16-QAM, respectively. The Monte-Carlo simulations using the BER vs. SNR performance metric are considered for the 2nd and 5th iterations in the case of convolutional BICM utilizing the RSC channel encoder with constraint length $K = 5$. The soft decoder employed in this paper, as mentioned previously, is a MAP decoder. In addition, different combinations of $\left(N_{tx}^a, N_{rx}^b, N_{tx}^b\right)$ are used, i.e. $(2, 4, 2)$, $(2, 6, 2)$, and $(2, 8, 2)$, where $\left(N_{tx}^a, N_{rx}^b\right)$ denotes the number of antennas being used between the transmitting port of node a and the receiving port of node b, respectively. This represents the path of the desired signal. Furthermore, N_{tx}^b denotes the number of antennas at the transmitting port of node b causing self-interference due the FD operation. We assume the number of transmit antennas is less than receive antennas for the sake of increasing the diversity gain and also to reduce the source of SI and transmit noise. This will enhance the overall SINR of the FD-MIMO system [3]. Moreover, a comparison is presented between the system performance under a residual (INR) of 10 dB against the interference-free case.

From these figures, it is clear that the performance under this INR after five iterations is very close to the SI free performance, which demonstrates that the proposed SIC has a near-optimal achievement, especially with higher number of antennas at the receiving port, N_{rx}^b, and for a fixed number of transmitting antennas, N_{tx}^a. It is worth noting that in the three investigated scenarios the performance improves as the number of iterations is increased. Furthermore, we

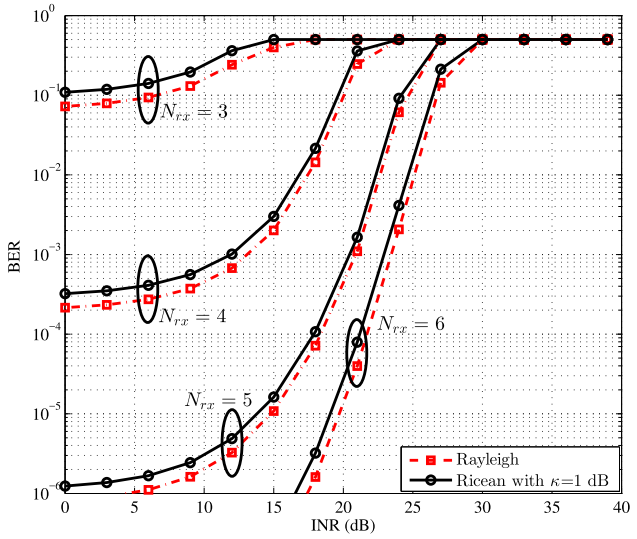

Fig. 8. BER vs. INR performance of FD-MIMO utilizing 16-QAM with 1/2-rate convolutional BICM for different combinations of $(2, N_{rx}, 2)$ and SNR = 10 dB.

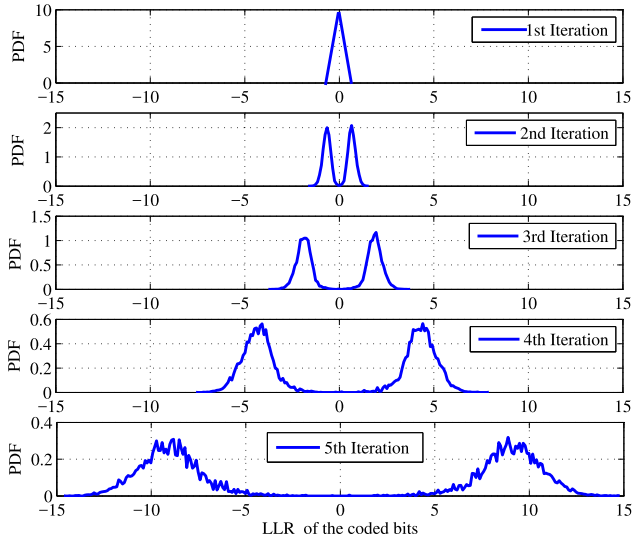

Fig. 9. PDF distributions vs. iterations for a $(2, 4, 2)$-FD-MIMO system at SNR = 5 dB and residual INR of 5 dB.

demonstrate in these two figures the derived tight upper bound on the performance of the proposed system under the same conditions described above, which reveals that at high SNR the tight upper bound performances match closely the results obtained from simulations of the proposed system after the 2nd iteration.

Fig. 8 demonstrates the BER vs. INR performance of the proposed system for the two sceneries of Rayleigh and low κ-factor distributions. Different combinations of $(2, N_{rx}, 2)$ are used, such as $N_{rx} = 4, 5, 6$, for rate-1/2 RSC codes with constraint length $K = 5$ as shown in Table II, and for the case of 16-QAM modulation scheme. Moreover, the SNR is fixed at 10 dB. This figure reveals that the performance is degraded with increasing the INR, but by increasing the number of the receive antennas for a particular number of transmit antennas, the system becomes more tolerant against the residual SI.

Fig. 9 shows the probability density function (PDF) distributions of the LLRs related to the coded bits, which represent the output of the soft decoder, for the 1st to 5th iterations of the $(2, 4, 2)$-MIMO system for the RSC convolutional code with $K = 5$ under residual INR = 5 dB and at an SNR = 5 dB. The PDF distribution of the LLRs during the first iteration is completely different from the expected bi-modal Gaussian PDF. As explained previously in Section II-E, this iteration performs the initialization of the IDD in order to start the second iteration with the required soft information. Note that as the number of iterations increases, the PDF converges to a bi-modal Gaussian PDF representing the logic (0, 1) bits.

VI. Conclusions

In this paper, an active SIC for a bi-directional coded FD-MIMO transceiver has been proposed and its performance has been evaluated. The proposed receiver utilizes IDD, which comprises soft-PIC and adaptive MMSE filtering. This SIC approach has been implemented by exchanging the soft information of the signal of interest and the SI signal between the equalizer and the soft decoder. The equalizer performs both adaptive MMSE filtering to remove residual interference and LLR demapping. Furthermore, soft channel decoding of the BICM has been used by employing a MAP decoder. The system performance was evaluated using numerical simulations obtained for several combinations of transmitting and receiving MIMO antennas in the presence of AWGN and over independent MIMO flat fading channels. Additionally, a tight upper bound on the performance of the proposed receiver is derived under the same conditions to validate the simulation results. The modulation scheme used in this paper was M-QAM. The obtained results demonstrate that with increasing number of iterations the FD-MIMO receiver can reconstruct the desired signal and the interference more precisely. Furthermore, for a given number of SI antennas, increasing the number of receive antennas enhances the tolerance to residual interference. The proposed coded FD-MIMO transceiver with IDD offers increased resilience to interference power, which in turn enhances the overall system performance. Furthermore, the EXIT chart based results demonstrating the convergence of the proposed iterative system reveal that the number of iterations required for the soft decoder and the equalizer to converge is inversely proportional to the number of the received antennas.

Appendix
Moment Generating Function of SINR

We start by evaluating the cumulative distribution function (CDF) of SINR, γ_l, in (23) which can be expressed as [41], [48]

$$F(\gamma_l) = 1 - e^{-\psi\gamma_l} \sum_{n=1}^{N_{rx}} \frac{A_n(\gamma_l)}{\Gamma(n)} (\psi\gamma_l)^{n-1}, \tag{47}$$

where $\psi = 1/\sigma^2_{\zeta_{n,l}}$, in which $\sigma^2_{\zeta_{n,l}}$ represents the variance of residual-SI-plus-noise at node b in the digital domain, as defined earlier in (21). $\Gamma(n) = (n-1)!$ represents the Gamma

function while $A_n(\gamma_l)$ represents an auxiliary function which is defined as

$$A_n(\gamma_l) = \begin{cases} 1, & \text{for } n \leq N_{div}, \\ \dfrac{1}{(\gamma_l+1)^{N_{tx}-1}} \displaystyle\sum_{i=0}^{N_{rx}-n} C_i \gamma^i, & \text{for } n > N_{div}, \end{cases} \quad (48)$$

where $N_{div} = (N_{rx} - N_{tx} + 1)$ represents the diversity order of the MIMO system, and C_i represents the coefficients of x^i in the term $(1+x)^{N_{tx}-1}$ for $i \geq 0$. For instance, if $N_{tx} = 2$ this term will appear as $(1+x)$, which means that $C_0 = 1$ and $C_1 = 1$. The MGF of γ_l as a function of the receive and transmit antennas, N_{rx} and N_{tx}, respectively, can be defined as

$$\mathcal{M}_{N_{tx},N_{rx}}(t) \triangleq \int_0^\infty e^{-t\gamma_l} dF(\gamma_l) = t\int_0^\infty e^{-t\gamma_l} F(\gamma_l) d\gamma_l, \quad (49)$$

which can be solved using the techniques presented in [40] and [43] resulting in

$$\begin{aligned} \mathcal{M}_{N_{tx},N_{rx}}(t) = &\left[\frac{\psi}{\psi+t}\right]^{N_{div}} - te^{(t+\psi)} \\ &\times \sum_{n=N_{div}+1}^{N_{rx}} \sum_{i=1}^{N_{rx}-n} \sum_{k=0}^{n+i-1} \frac{\psi^{n-1} C_i D_k}{\Gamma(n)} \\ &\times E_{N_{tx}-k-1}(\psi+t), \end{aligned} \quad (50)$$

where D_k represents the coefficients of x^k in the term $(x-1)^{n+i-1}$, which can be computed as $D_k = (-1)^k \binom{n}{k}$, for $k = 0, 1, 2, \ldots, n$. Additionally, $E_n(\alpha) \triangleq \int_1^\infty \frac{e^{-\alpha t}}{t^n} dt, \forall \alpha > 0$ is the generalized exponential integral function.

ACKNOWLEDGMENT

No new data were created during this study.

REFERENCES

[1] 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.

[2] Z. Zhang, X. Chai, K. Long, A. V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: Self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128–137, May 2015.

[3] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983–5993, Dec. 2011.

[4] T. Riihonen, S. Werner, and R. Wichman, "Spatial loop interference suppression in full-duplex MIMO relays," in *Proc. 43rd Asilomar Conf. Signals. Syst.*, Nov. 2009, pp. 1508–1512.

[5] H. Ju, E. Oh, and D. Hong, "Improving efficiency of resource usage in two-hop full duplex relay systems based on resource sharing and interference cancellation," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 3933–3938, Aug. 2009.

[6] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 2, pp. 913–927, Feb. 2014.

[7] H. Q. Ngo, H. A. Suraweera, M. Matthaiou, and E. G. Larsson, "Multipair full-duplex relaying with massive arrays and linear processing," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1721–1737, Sep. 2014.

[8] M. A. Ahmed and C. C. Tsimenidis, "Coded full-duplex MIMO with iterative detection and decoding," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2015, pp. 4859–4864.

[9] M. A. Ahmed, C. C. Tsimenidis, and A. A. Rawi, "Performance analysis of full-duplex-MRC-MIMO with self-interference cancellation using null-space-projection," *IEEE Trans. Signal Process.*, vol. 64, no. 12, pp. 3093–3105, Jun. 2016.

[10] M. A. Ahmed and C. C. Tsimenidis, "A tight upper bound on the performance of iterative detection and decoding for coded full-duplex SIMO systems," *IEEE Commun. Lett.*, vol. 20, no. 3, pp. 606–609, Mar. 2016.

[11] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1541–1553, Sep. 2012.

[12] B. P. Day, D. W. Bliss, A. R. Margetts, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," in *Proc. 45th Asilomar Conf. Signals. Syst.*, Nov. 2011, pp. 1386–1390.

[13] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Proc. 44th Asilomar Conf. Signals. Syst. Comput.*, Nov. 2010, pp. 1558–1562.

[14] A. M. Tonello, "Space-time bit-interleaved coded modulation with an iterative decoding strategy," in *Proc. IEEE Veh. Technol. Conf.*, vol. 1. Sep. 2000, pp. 473–478.

[15] S. Sadough and M.-A. Khalighi, "Optimal turbo-blast detection of MIMO-OFDM systems with imperfect channel estimation," in *Proc. IEEE 18th Int. Symp. Pers., Indoor Mobile Radio Commun. (PIMRC)*, Sep. 2007, pp. 1–6.

[16] N. Du and D. Xu, "Low complexity iterative receiver for MIMO-OFDM systems," in *Proc. Int. Symp. Intel. Signal Process. Commun. Syst. (SPACS)*, Nov. 2007, pp. 622–625.

[17] J. Y. Hwang, D.-K. Cho, K. S. Kim, and K.-C. Whang, "A Turbo-BLAST method with non-linear MMSE detector for MIMO-OFDM systems," in *Proc. IEEE Asia Pacific Conf. Circuits Syst. (APCCAS)*, Dec. 2006, pp. 295–297.

[18] H. Lee, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, vol. 12. Jun. 2006, pp. 5377–5382.

[19] M. Sellathurai and S. Haykin, "Turbo-BLAST for wireless communications: Theory and experiments," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2538–2546, Oct. 2002.

[20] M. Sellathurai and S. Haykin, "Turbo-BLAST: Performance evaluation in correlated Rayleigh-fading environment," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 340–349, Apr. 2003.

[21] D. Na, G. Pin-biao, and C. Ning, "An improved iterative receiver scheme for turbo-BLAST system with anti-gray mapping," in *Proc. 3rd Int. Symp. Inform. Process. (ISIP)*, Oct. 2010, pp. 411–414.

[22] D. Na, G. Pin-biao, and C. Ning, "A low-complexity iterative receiver scheme for turbo-BLAST system," in *Proc. IEEE 10th Int. Conf. Signal Process. (ICSP)*, Oct. 2010, pp. 1548–1551.

[23] T. Ait-Idir, S. Saoudi, and N. Naja, "Space–time turbo equalization with successive interference cancellation for frequency-selective MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 57, no. 5, pp. 2766–2778, Sep. 2008.

[24] T. Abe and T. Matsumoto, "Space-time turbo equalization in frequency-selective MIMO channels," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 469–475, May 2003.

[25] T. Abe, S. Tomisato, and T. Matsumoto, "A MIMO turbo equalizer for frequency-selective channels with unknown interference," *IEEE Trans. Veh. Technol.*, vol. 52, no. 3, pp. 476–482, May 2003.

[26] T. Riihonen and R. Wichman, "Analog and digital self-interference cancellation in full-duplex MIMO-OFDM transceivers with limited resolution in A/D conversion," in *Proc. 46th Asilomar Conf. Signals. Syst. Comput.*, Nov. 2012, pp. 45–49.

[27] D. W. Bliss and Y. Rong, "Effects of channel estimation errors on in-band full-duplex MIMO radios using adaptive transmit spatial mitigation," in *Proc. Asilomar Conf. Signals. Syst Comput.*, Nov. 2013, pp. 9–13.

[28] J. Hu and T. M. Duman, "Low density parity check codes over wireless relay channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 9, pp. 3384–3394, Sep. 2007.

[29] A. Sadeghi, F. Lahouti, and M. Zorzi, "Constellation shaping and LDPC coding in a bidirectional full duplex communication," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*. Dec. 2015, pp. 1–6.

[30] R. A. Carrasco and M. Johnston, *Non-Binary Error Control Coding for Wireless Communication Data Storage*. Hoboken, NJ, USA: Wiley, 2008.

[31] X. Xiong, X. Wang, T. Riihonen, and X. You, "Channel estimation for full-duplex relay systems with large-scale antenna arrays," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6925–6938, Oct. 2016.

[32] H. Ju, S. Lee, K. Kwak, E. Oh, and D. Hong, "A new duplex without loss of data rate and utilizing selection diversity," in *Proc. IEEE Veh. Technol. Conf. (VTC)*. May 2008, pp. 1519–1523.

[33] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296–4307, Dec. 2012.
[34] T. Chen and S. Liu, "A multi-stage self-interference canceller for full-duplex wireless communications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.
[35] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," *IEEE Trans. Wireless Commun.*, vol. 13, no. 7, pp. 3821–3836, Jul. 2014.
[36] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proc. 11th USENIX Symp. Netw. Syst. Design Implement. (NSDI)*, 2014, pp. 359–372.
[37] L. Fang, Q. Guo, D. Huang, and S. Nordholm, "A low cost soft mapper for turbo equalization with high order modulation," in *Proc. Int. SoC Design Conf. (ISOCC)*, Nov. 2012, pp. 305–308.
[38] A. Tomasoni, M. Ferrari, D. Gatti, F. Osnato, and S. Bellini, "A low complexity turbo MMSE receiver for W-LAN MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, vol. 9. Jun. 2006, pp. 4119–4124.
[39] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.
[40] N. Kim and H. Park, "Bit error performance of convolutional coded MIMO system with linear MMSE receiver," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3420–3424, Jul. 2009.
[41] J. Proakis and M. Salehi, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2008.
[42] P. Frenger, P. Orten, and T. Ottosson, "Convolutional codes with optimum distance spectrum," *IEEE Commun. Lett.*, vol. 3, no. 11, pp. 317–319, Nov. 1999.
[43] M. R. McKay and I. B. Collings, "Error performance of MIMO-BICM with zero-forcing receivers in spatially-correlated Rayleigh channels," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 787–792, Mar. 2007.
[44] R. Durrett, *Probability: Theory Examples*. Cambridge, MA, USA: Cambridge Univ. Press, 2010.
[45] S. Ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct. 2001.
[46] M. El-Hajjar and L. Hanzo, "EXIT charts for system design and analysis," *IEEE Commun. Surv. Tuts.*, vol. 16, no. 1, pp. 127–153, 1st Quart., 2014.
[47] S. Ten Brink, "Designing iterative decoding schemes with the extrinsic information transfer chart," *AEU Int. J. Electron. Commun.*, vol. 54, no. 6, pp. 389–398, 2000.
[48] A. Hedayat and A. Nosratinia, "Outage and diversity of linear receivers in flat-fading MIMO channels," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5868–5873, Dec. 2007.

Mohamad A. Ahmed (S'13–M'17) received the B.Sc. degree in electrical engineering (electronic and communication) from Mosul University, Mosul, Iraq, in 1999, and the M.Sc. degree (Hons.) and the Ph.D. degree in communication and signal processing from the School of Electrical and Electronics Engineering, Newcastle University, Newcastle Upon Tyne, U.K., in 2009 and 2017, respectively. He is currently a Lecturer in signal processing for communications with the College of Electronics Engineering, Ninavah University, Mosul. His research emphasizes the field of full-duplex-MIMO wireless communications. His current research interests include numerical, simulation, and theoretical performance analyses for the next generation of wireless communication (5G). Moreover, he is interested in the subjects of wireless communications, microwave, radio frequency engineering, and digital signal processing.

Charalampos C. Tsimenidis (M'05–SM'12) received the M.Sc. degree (Hons.) and the Ph.D. degree in communications and signal processing from Newcastle University in 1999 and 2002, respectively. He is currently a Reader in digital communications with the School of Engineering, Newcastle University. During the last 15 years, he has published over 180 conference and journal papers, supervised successfully 3 M.Phil. and 36 Ph.D. students, and made contributions in the area of receiver design to several European and U.K. funded research projects. His main research interests are in the area of adaptive array receivers for wireless communications including demodulation and error correction algorithms for doubly spread multi-path channels. He is a member of the IET.