

Improper Gaussian Signaling for Hardware Impaired Multihop Full-Duplex Relaying Systems

Sidrah Javed, Osama Amin, Basem Shihada and Mohamed-Slim Alouini

Abstract

In this paper, we analyze the performance degradation of a multi-hop decode-and-forward full-duplex relaying system (MH-DF-FDR) caused by the residual self-interference (RSI) and hardware distortions (HWD) imposed by the FDR operation and imperfect hardware, respectively. In addition, we study the benefits of employing improper Gaussian signaling (IGS) in the MH-FDR system. Different from the traditional symmetric signaling scheme, i.e., proper Gaussian signaling (PGS), IGS has non-zero pseudo-variance that can limit the impact of RSI and HWD on the MH-FDR system. To evaluate the system performance gain using IGS, first we express the end-to-end achievable rate of the MH system as the minimum rate supported by all participating links. Then, we optimize the pseudo-variance of all participating transmitters including source and relays to compensate the interference impact and improve the end-to-end achievable rate. We propose two network optimization schemes based on the system characteristics i.e. joint optimization framework and distributed optimization scenario. Interestingly, IGS-based scheme outperforms its counterpart PGS-based scheme, especially at higher interference-to-noise ratio. Our findings reveal that using IGS in single-user detection systems that suffer from both RSI and HWD can effectively mitigate the degradation in the achievable rate performance.

Index Terms

Achievable rate, convex-concave procedure, full-duplex relaying, hardware distortions, improper Gaussian signaling, residual self-interference, and sequential convex programming.

S. Javed, O. Amin, B. Shihada and M.S. Alouini are with CEMSE Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia.

E-mail: {sidrah.javed, osama.amin, basem.shihada, slim.alouini}@kaust.edu.sa

I. INTRODUCTION

The ever-increasing demand of ubiquitous high data rates, with low latency, extended coverage, high energy efficiency and significantly higher bandwidth per subscriber, are the major driving forces for the upcoming wireless communications. Several studies have been carried out to investigate extreme node densification and collaborative radio technologies to improve the spectral efficiency and meet the exponentially growing wireless data traffic demands [1], [2]. Multi-hop relaying technology has gained much interest as it can extend the coverage and improve the energy efficiency [3], [4]. However, the spectral efficiency decreases with the increase of number of relays/hops, where the frequency (or time) is shared between the nodes to deliver the data to the destination [5]. FDR is a promising cooperative technology that can compensate the spectral efficiency loss by allowing each node to transmit and receive simultaneously [6]. FDR can be used in different network topologies and applications to improve the quality-of-service such as unmanned autonomous vehicles and self-driving robots [7].

Although FDR systems allow better resource exploitation than half-duplex relaying systems by simultaneously transmitting and receiving on the same channel, they suffer from some challenges that can limit their operation. Self-interference (SI) is one of the main deterrents in employing such systems. The SI signal is relatively larger than the desired signal of interest, which increases the dynamic range span of the low-noise amplifier (LNA) and the analog-to-digital converter (ADC) at the receiver side. As a result, both the undesired interference and the hardware noise levels increase, which can greatly suppress or even destroy the information bearing signal [8]. The SI of FDR systems can be canceled through multiple analog and digital stages to ease the detection of the desired signal [9], [10]. However, the perfect knowledge of the pilot signal is not sufficient to cancel the SI completely due to the channel estimation errors and the limited dynamic range of the filters, amplifiers and ADC. Therefore, the FDR performance is limited by the RSI after different cancellation stages. Furthermore, MH-FDR performance can be further degraded due to the aggregate effect of RSI from different relaying stages.

HWD is another deteriorating factor that affects communication systems performance generally and MH-FDR systems particularly. HWD is caused by different circuit impairments such as the phase noise, non-linear distortion, and in-phase and quadrature imbalance (IQI) [11]. The performance of MH-FDR can be severely affected by the HWD due to the accumulated effect of impaired transceivers at the relaying nodes.

A. *Related Work:*

Various studies have been carried out to address the deteriorating impact of RSI or/and HWD in full-duplex communication systems. For example, the effects of hardware impairments were investigated for different full-duplex communication systems in [12]–[19] and particularly for relay systems in [14]–[19]. Of all these contributions, [12] proposes a digital cancellation scheme to supplement RF/analog cancellation techniques for self-interference mitigation in single-channel FD wireless communication. On the other hand, [19] focuses on a low complexity hardware impairments aware transceiver scheme to mitigate distortions in the transmitter and the receiver. Similarly, [15], [16], [18] emphasis on the system performance analysis under IQI at relay and destination with ideal transmitter considering a half-duplex amplify-and-forward (AF) relay and orthogonal frequency-division multiplexing (OFDM) modulation, and IQI at all nodes in OFDM dual-hop opportunistic AF relaying. Similarly, other contributions [20]–[24] presented power control mechanisms to improve system throughput or outage performance in the presence of RSI and inter-relay interference assuming ideal transceivers.

However, very few studies focus on both RSI and HWD in a full-duplex operation mode. For example, [14] carries out the performance analysis of dual-hop proactive DF relaying networks with best relay selection under hardware impairment and co-channel interference. Similarly, [17] analyzes the outage probability of dual-hop DF FDR for an OFDM system in the presence of IQI and loop-back SI. Besides carrying out the performance analysis, few works proposed some compensation schemes to improve the system performance. For example, [13] proposes a novel widely linear digital cancellation processing to mutually mitigate SI and practical hardware imperfections in direct-conversion FD transceiver. In [25] and [26], authors proposed compensation schemes for various HWIs and RSI at the AF-relay, considering dual-hop HWI-FD-AF-relay system with ideal source and destination. This motivates us to propose a mitigation signaling design to concurrently combat RSI and transceiver distortions at the source and destination besides multiple HWI-FD-relays. Furthermore, we have focused on the DF relaying strategy in-place of AF relaying scheme, to support the communication in an interference limited environment, in order to meet next generation traffic demands.

Analyzing the impact of hardware imperfections and RSI on the system performance and evaluating different compensation schemes require an accurate statistical model of these imperfections. RSI and HWD are modelled by widely linear transformations as discussed in [13]

and [27], respectively. Thus, according to statistical signal processing studies, widely linear precoders/transformations can efficiently map symmetric information-bearing signals to asymmetric signals at each transmitter [28], [29]. Therefore, this work models both RSI and HWD as asymmetric signals. Furthermore, the proposed research can employ IGS for signal transmission to jointly mitigate the deterrent effect of both RSI and HWD. IGS scheme has already been proven to evidently improve system performance in various system configurations such as multiple-input multiple output systems [30], [31], cognitive radio systems [32], [33], full-duplex relaying [27], [34], [35], and alternating relaying [36]. The IGS transmission scheme is expected to outperform PGS in the presence of RSI and HWD in a MH-DF-HWD-FDR system. The aim is to quantify the gain obtained by optimal IGS over PGS and to evaluate if the gain is significant enough to adopt IGS optimization framework as the optimal IGS solution can sometimes reduce to PGS. In order to maximize the end-to-end rate, the IGS transmit signaling characteristics of the source and relay have already been optimized in the conference version of this paper for dual-hop DF-FDR system [37].

B. Paper Contribution:

In this paper, we study the utilization of asymmetric signaling scheme instead of the symmetric signaling scheme to combat both the RSI and HWD in MH-DF-FDR systems. Symmetric signaling or PGS is the traditional signaling scheme that assumes independent signal components with equal power, which is described by its variance. On the other hand, asymmetric signaling or IGS relaxes the PGS characteristics and can have dependent signal component with/without equal power. Therefore, an IGS needs an additional statistical quantity to be accurately characterized, which is called the pseudo-variance [38]. We should note that IGS can be practically implemented using widely linear precoders, which efficiently maps symmetric information-bearing signals to asymmetric signals at each transmitter [28], [29]. The main contributions of this paper are summarized as follows:

- Employment of multiple DF-FDRs to establish a communication link between two distant nodes in the absence of a direct link assuming imperfect hardware transceivers.
- Studying the causes of the MH-FDR system performance degradation, i.e. RSI and HWD, based on a realistic asymmetric statistical model relative to [29] which studies the multi-user interference in a two-user SISO interference channel model.

- Investigating the end-to-end achievable rate performance of the adopted MH-DF-HWD-FDR system with asymmetric signals and interference sources.
- Developing a rigorous joint optimization framework to design the transmission parameters in order to maximize the overall data-rate and effectively mitigate the degradation effects of the undesired interfering signals. Joint optimization of the transmit parameters of all participating relays and source transmitter has been carried out based on an different optimization framework relative to [29] and [37]. For instance, [29] employed Gaussian Randomization Procedure for transmit optimization and [37] proposed a two-loop optimization framework employing generalized Dinkelbach algorithm and successive convex programming, respectively. On the other hand, the present scheme merely relies on successive convex programming reducing the computational complexity and processing time.
- Developing a distributed optimization framework that suits practical implementation of the proposed transmission scheme offering reduced round-trip delays, computational complexity and communication overhead.

C. Organization and Notation:

The rest of the paper is organized in the following fashion. Section II studies the statistical model for the MH-DF-FDR system under asymmetric self-interference and transceiver distortions. Section III focuses on the information-theoretic achievable rate in the presence of improper Gaussian interference and signal. Section IV focuses on designing the transmission parameters based on the RSI and HWD characteristics. It deals with the joint as well as distributed optimization framework to fine-tune the statistical IGS parameters to achieve optimum system performance. Section V numerically analyzes the system performance with/without IGS assuming different system parameters.

Notations: In this paper, scalars are denoted by lower-case italic letters, while vectors and matrices are denoted by boldfaced lower-and upper-case letters, respectively. For a complex scalar x , its conjugate and absolute value are represented by x^* and $|x|$, respectively. Moreover, $\Re\{x\}$ and $\Im\{x\}$ represent the real and imaginary components of a complex scalar, respectively. The expected value operator and the floor operator to round the rational number to the greatest lower integer are given by $\mathbb{E}[\cdot]$ and $\lfloor \cdot \rfloor$, respectively. On the other hand, for a given vector \mathbf{x} , its complex-conjugate, transpose and conjugate-transpose are represented by \mathbf{x}^* , \mathbf{x}^T and \mathbf{x}^H , respectively. In addition, $\text{diag}\{\mathbf{x}\}$ operator represents a diagonal matrix whose diagonal entries

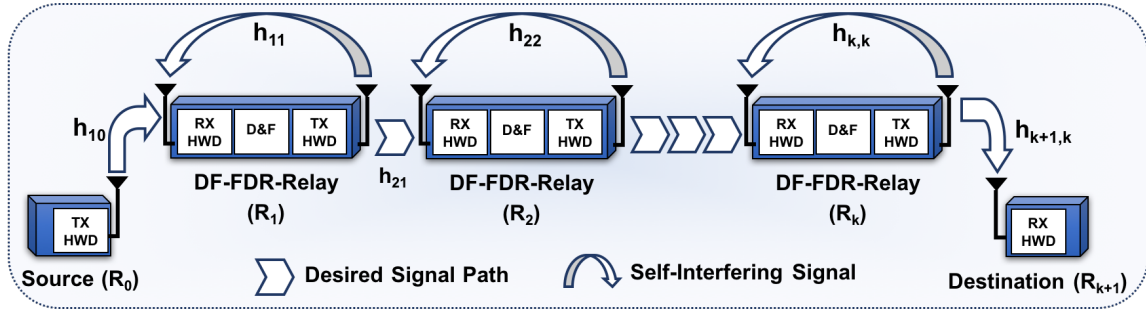


Fig. 1. FDR system under HWD and RSI.

are the elements of vector \mathbf{x} and $\nabla \mathbf{x}$ illustrates the gradient of a vector \mathbf{x} . The right circular shift operator of a vector \mathbf{x} for n shifts is defined as $\langle \mathbf{x} \rangle_{(RS,n)}$, e.g., $\langle [a \ b \ c \ d]^T \rangle_{(RS,1)} = [d \ a \ b \ c]^T$. Moreover, to append m zeros at the end of a vector \mathbf{x} we use the following notation, $\langle \mathbf{x} \rangle_{(AZ,m)}$, e.g., $\langle [a \ b \ c]^T \rangle_{(AZ,2)} = [a \ b \ c \ 0 \ 0]^T$.

II. SYSTEM DESCRIPTION

Consider a MH relaying system, where a source (R_0) intends to communicate with a destination (R_{k+1}) as shown in Fig. 1. Both the high shadowing and the severe path loss effect are responsible for the absence of the direct link between R_0 and R_{k+1} . As such, multiple full duplex relays ($R_1 - R_k$) operate as intermediate nodes to facilitate the end-to-end communication by expanding the coverage area with a full transmission rate. The limited power budget of the participating relays renders negligible inter-relay interference. All relays operate in a DF relaying strategy and the SI at the relays can only be canceled partially, thus resulting in RSI as depicted in Fig. 1. Furthermore, various HWDs at the transmitter and receiver RF branches can drastically degrade the overall system performance. Before describing the HWD model and FDR system model, we define the adopted statistical signal model in the following subsection.

A. Statistical Signal Model

To characterize the difference between symmetric and asymmetric signals, we consider a complex Gaussian random variable x and introduce the following definitions:

Definition 1. [38] *The variance and the pseudo-variance of x are defined, respectively, as $\sigma_x^2 = \mathbb{E}[|x|^2]$ and $\tilde{\sigma}_x^2 = \mathbb{E}[x^2]$.*

Definition 2. [38] A complex random variable is called proper if its pseudo-variance is equal to zero, otherwise it is called improper .

Definition 3. Circularity coefficient \mathcal{C}_x of a random variable x is defined as the ratio of the absolute pseudo-variance of x and its variance, i.e., $\mathcal{C}_x = |\tilde{\sigma}_x^2|/\sigma_x^2$, where $0 \leq \mathcal{C}_x \leq 1$ [39].

The circularity coefficient measures the degree of impropriety or symmetry of x , where $\mathcal{C}_x = 0$ indicates proper or symmetric signal and $\mathcal{C}_x = 1$ indicates maximally improper or asymmetric signal.

Definition 4. The complex Gaussian random variable x is fully described as: $x \sim \mathcal{CN}(m_x, \sigma_x^2, \tilde{\sigma}_x^2)$, where m_x is the statistical average of x , i.e., $m_x = \mathbb{E}[x]$.

B. Distortion Model

In this subsection, we start by describing the mathematical model of the aggregated HWD in radio frequency transceivers for a single link system. Then, we extend the model to capture the generalized scenario of MH-FDR system. A transmit signal x undergoes transceiver distortions as well as fading channel before being received at the destination as

$$y = \sqrt{p}h(x + \eta_{\text{tx}}) + \eta_{\text{rx}} + z, \quad (1)$$

where p is the transmitted power, h is the fading channel, and z is the additive white Gaussian noise (AWGN) with variance σ_z^2 . Moreover, η_{tx} and η_{rx} are the additive impairment distortions at the transmitter and the receiver, respectively. Various theoretical investigations and measurement results indicate that the Gaussian model accurately describes the aggregate of all residual RF impairments when compensation algorithms are applied to mitigate hardware impairments ([14], [19], [40]–[49] and references therein). This can also be motivated analytically by the central limit theorem. In addition, η_{tx} and η_{rx} are generalized as asymmetric signals pertaining to the transformation caused by some hardware impairments such as HWD with wide linear transformation characteristics [28], [29]. Thus, the aggregate HWDs at the transmitter and the receiver are random variables with $\eta_{\text{tx}} \sim \mathcal{CN}(0, \kappa_{\text{tx}}, \tilde{\kappa}_{\text{tx}})$ where $|\tilde{\kappa}_{\text{tx}}| \leq \kappa_{\text{tx}} \leq \sigma_x^2$ and $\eta_{\text{rx}} \sim \mathcal{CN}(0, p|h|^2\kappa_{\text{rx}}, ph^2\tilde{\kappa}_{\text{rx}})$, where $|\tilde{\kappa}_{\text{rx}}| \leq \kappa_{\text{rx}} \leq \sigma_x^2$. Clearly, the proposed model reduces to the well-known system model $y = \sqrt{p}hx + z$ in the absence of transceiver distortions i.e. $\eta_{\text{tx}} = \eta_{\text{rx}} = 0$.

Lemma 1. *The equivalent generalized aggregate model of the HWD is given by*

$$y = \sqrt{p}h(x + \eta) + z, \quad (2)$$

with $\eta = \eta_{\text{tx}} + \eta_{\text{rx}}$ is distributed as $\mathcal{CN}(0, \kappa, \tilde{\kappa})$. Moreover, $\kappa = \kappa_{\text{tx}} + \kappa_{\text{rx}}$ and $\tilde{\kappa} = \tilde{\kappa}_{\text{tx}} + \tilde{\kappa}_{\text{rx}}$ capture the aggregate HWD at both the transmitter and the receiver along with impact of fading channels.

Proof. We assumed a general asymmetric model for the additive distortion with $\eta_{\text{tx}} \sim \mathcal{CN}(0, \kappa_{\text{tx}}, \tilde{\kappa}_{\text{tx}})$ and $\eta_{\text{rx}} \sim \mathcal{CN}(0, p|h|^2\kappa_{\text{rx}}, ph^2\tilde{\kappa}_{\text{rx}})$. Different from the existing literature that assumes symmetric HWD, we assume the general asymmetric scenario where having symmetric distortion at both in-phase and quadrature components is not the only possible scenario. Moreover, the symmetric distortion can be transformed into asymmetric one after passing through widely linear transformation [28]. For a given fading channel, the variance and pseudo variance of the aggregated impairments in (1) are expressed, respectively, as follows

$$\mathbb{E} \left[|\sqrt{p}h\eta_{\text{tx}} + \eta_{\text{rx}}|^2 \right] = p|h|^2 (\kappa_{\text{tx}} + \kappa_{\text{rx}}), \quad (3)$$

$$\mathbb{E} \left[(\sqrt{p}h\eta_{\text{tx}} + \eta_{\text{rx}})^2 \right] = ph^2 (\tilde{\kappa}_{\text{tx}} + \tilde{\kappa}_{\text{rx}}). \quad (4)$$

Moreover, the variance and pseudo-variance of the aggregated asymmetric distortions in (2) are $p|h|^2\kappa$ and $ph^2\tilde{\kappa}$, respectively. Thus, the signal model in (1) can be equivalently modeled as in (2) when $\kappa = \kappa_{\text{tx}} + \kappa_{\text{rx}}$ and $\tilde{\kappa} = \tilde{\kappa}_{\text{tx}} + \tilde{\kappa}_{\text{rx}}$. \square

It is important to note that in case of ideal hardware i.e. $\eta = 0$ is imposed by $\kappa = 0$ and $\tilde{\kappa} = 0$. Where, $\kappa = 0$ is dictated by the negligible transmitter and receiver distortion variances, $\kappa_{\text{tx}} = 0$ and $\kappa_{\text{rx}} = 0$, respectively. Moreover, $\tilde{\kappa} = 0$ follows from Definition 3.

C. FDR under HWD System Model

In the proposed MH-FDR with the DF relaying strategy under HWD, the m^{th} node R_m transmits an IGS signal $x_m \sim \mathcal{CN}(0, \sigma_m^2, \tilde{\sigma}_m^2)$ to the n^{th} node R_n in one hop. Pertaining to the FDR operation, the received signal at the relay node R_n suffers from RSI h_{nn} in addition to the aggregate effect of transceiver distortions for both m - n link, η_{nm} , and n - n link, η_{nn} . The signal is transmitted from R_0 to R_{k+1} in a sequential order. Therefore, the generalized received signal at the n^{th} receiver can be expressed as

$$y_n = \sqrt{p_m}h_{nm}(x_m + \eta_{nm}) + \sqrt{p_n}h_{nn}(x_n + \eta_{nn}) + z_n, \quad (5)$$

where $n = 1, 2, \dots, k + 1$ denotes the receiver nodes and $m = n - 1$ represents the sequential transmission nodes, thus $m = 0, 1, \dots, k$. In addition, p_m is the transmit power of the m^{th} node taken from a limited power budget, h_{nm} is the flat-fading channel of the m - n link¹ and $z_n \sim \mathcal{CN}(0, \sigma_z^2, 0)$ is independent identically distributed AWGN at the n^{th} receiver node. The receiver at the relay node decodes the transmitted signal using single user decoder, then encodes it from IGS code-book as $x_n \sim \mathcal{CN}(0, \sigma_n^2, \tilde{\sigma}_n^2)$ for further transmission. The same transmitted signal causes self-interference in the FDR transmission mode through the h_{nn} link. Measurement-driven experimental studies [9], [50], [51] have shown that, after undergoing all possible isolation/cancellation techniques and assuming the perfect cancellation of slowly-varying line-of-sight path, the residual interference can be well characterized as a flat-fading channel [17], [35], [52]–[55] and the RSI channel h_{nn} can be modeled as a proper complex Gaussian random variable with zero mean and π_{nn} variance ([56]–[60], and references therein).² It is worthy noting that the self-interference link does not exist at the destination node R_{k+1} as there is no further transmission and self-interference. Moreover, the transceiver HWD of the m - n link, i.e., η_{nm} , is assumed to demonstrate the following statistical characteristics: $\eta_{nm} \sim \mathcal{CN}(0, \kappa_{nm}, \tilde{\kappa}_{nm})$.

III. ACHIEVABLE RATES

The overall end-to-end achievable rate of the MH-DF-FDR system, R_T , is given as

$$R_T = \min_n \{R_{nm}\}; \quad m = n - 1, \quad (6)$$

where R_{nm} is the achievable rate of the R_m - R_n link in bits/sec. In our work, we deal with the RSI and HWD as interference terms, thus R_{nm} considering the IGS transmission scheme and asymmetric HWD terms can be obtained as [29]

$$R_{nm} = \frac{B}{2} \log_2 \frac{\sigma_{y_n}^4 - |\tilde{\sigma}_{y_n}^2|^2}{\sigma_{I_n}^4 - |\tilde{\sigma}_{I_n}^2|^2}, \quad (7)$$

where B is the bandwidth of the channel. Moreover, $\sigma_{y_n}^2$ and $\tilde{\sigma}_{y_n}^2$ are the variance and the pseudo-variance of the received signal at the n^{th} node, respectively. Additionally, $\sigma_{I_n}^2$ and $\tilde{\sigma}_{I_n}^2$ are the

¹It is important to mention that the narrow band assumption has been adopted for the simplified setup. However, the same contribution can be straight-forwardly extended to the multipath channel scenario such as OFDM for each sub-channel / sub-band after incorporating the effective inter-carrier interference. The extension to OFDM does not affect the optimization framework as the various transmission streams do not share the common resource budget.

²Note that these assumptions do not compromise the insights of the analysis that follows. The same conclusions can be reached if more complex models are used.

variance and the pseudo-variance of the self-interference signal plus noise at the receiver end. Therefore, the achievable rate of the link between m^{th} transmit node and n^{th} receiver node, R_{nm} reduces to

$$R_{nm}(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2) = \frac{B}{2} \log_2 \frac{\alpha_{nm} - |p_m h_{nm}^2 (\tilde{\sigma}_m^2 + \tilde{\kappa}_{nm}) + p_n h_{nn}^2 (\tilde{\sigma}_n^2 + \tilde{\kappa}_{nn})|^2}{\beta_{nm} - |p_m h_{nm}^2 \tilde{\kappa}_{nm} + p_n h_{nn}^2 (\tilde{\sigma}_n^2 + \tilde{\kappa}_{nn})|^2}, \quad (8)$$

where β_{nm} and α_{nm} are defined, respectively as,

$$\beta_{nm} = (p_m |h_{nm}|^2 \kappa_{nm} + p_n |h_{nn}|^2 (\sigma_n^2 + \kappa_{nn}) + \sigma_z^2)^2, \quad (9)$$

$$\alpha_{nm} = (p_m |h_{nm}|^2 (\sigma_m^2 + \kappa_{nm}) + p_n |h_{nn}|^2 (\sigma_n^2 + \kappa_{nn}) + \sigma_z^2)^2. \quad (10)$$

According to (8), R_{nm} is a function of the pseudo-variance $\tilde{\sigma}_m^2$ of the transmitted signal and the pseudo-variance $\tilde{\sigma}_n^2$ of the self-interfering signal, which provides additional degrees of freedom to mitigate the asymmetric interference caused by the HWD as well as RSI. However, the achievable rate of the last hop between node R_k and R_{k+1} is only a function of $\tilde{\sigma}_k^2$ due to the absence of the self-interfering link.

IV. HWD- AND RSI-AWARE SIGNALING DESIGN

In this section, we design the transmit signals for all transmitting nodes, to maximize R_T under HWD and RSI for the adopted MH-DF-FDR system. The main goal of the system design is to optimize the statistical asymmetric characteristics of the transmitted signals to maximize R_T in (6) as follows

$$\begin{aligned} \mathbf{P1} : \quad & \text{maximize} \quad \min_n \{R_{nm}(\tilde{\sigma}_m^2, \tilde{\sigma}_n^2)\} \\ & \text{subject to} \quad 0 \leq |\tilde{\sigma}_m^2| \leq \sigma_m^2, \quad \forall m, \end{aligned} \quad (11)$$

where the constraint is adopted, to confirm the bounds on circularity coefficient, for the pseudo-variances of all transmitting signals from nodes R_0 to R_k . Throughout the rest of this work, we solve **P1** by proposing two optimization frameworks with different implementation, complexity and performance. Firstly, we propose an efficient joint optimization scheme which requires a centralized processing framework. Later, we present a distributed framework which optimizes the transmission parameters for a cluster of nodes.

A. Joint Optimization

In the joint optimization setup, we assume having a central node that gathers the channel state information (CSI), HWD and RSI levels from all the participating nodes to jointly optimize their signal parameters in order to maximize the overall end-to-end achievable rate. After processing the gathered information, the central node distributes the optimal transmission parameters for data transmission.

To solve **P1**, we write all optimization parameters (transmit pseudo-variances) in a vector form. The total number of the optimization parameters depend on the number of participating relays and the corresponding communication hops. Considering k intermediate relays between the source and the destination results in $k + 1$ hops and hence $k + 1$ transmitting nodes. Therefore, we need to optimize $k + 1$ complex pseudo-variance variables or $2(k + 1)$ real transmit pseudo-variance variables. Thus, we define a real vector \mathbf{s} that captures the real and imaginary variables as

$$\mathbf{s} = \left[\Re \{ \tilde{\sigma}_0^2 \} \quad \Im \{ \tilde{\sigma}_0^2 \} \quad \Re \{ \tilde{\sigma}_1^2 \} \quad \Im \{ \tilde{\sigma}_1^2 \} \quad \dots \quad \Re \{ \tilde{\sigma}_k^2 \} \quad \Im \{ \tilde{\sigma}_k^2 \} \right]^T. \quad (12)$$

Then, we express the link rate between nodes $R_m - R_n$ in (8) as a function of the vector \mathbf{s} carrying optimization variables:

$$R_{nm}(\mathbf{s}) = \frac{B}{2} \log_2 \frac{\alpha_{nm} - |\mathbf{s}^T \mathbf{u}_{nm} + v_{nm}|^2}{\beta_{nm} - |\mathbf{s}^T \mathbf{w}_{nm} + v_{nm}|^2}, \quad (13)$$

where \mathbf{u}_{nm} and \mathbf{w}_{nm} are defined, respectively, as follows,

$$\mathbf{u}_{nm} = \langle \langle [p_m h_{nm}^2 \quad j p_m h_{nm}^2 \quad p_n h_{nn}^2 \quad j p_n h_{nn}^2]^T \rangle_{(Az, 2k-2)} \rangle_{(Rs, 2m)}, \quad (14)$$

$$\mathbf{w}_{nm} = \langle \langle [p_n h_{nn}^2 \quad j p_n h_{nn}^2]^T \rangle_{(Az, 2k)} \rangle_{(Rs, 2n)}. \quad (15)$$

Moreover, the complex scalar v_{nm} is defined as a function of transmit power, CSI and HWD as,

$$v_{nm} = p_m h_{nm}^2 \tilde{\kappa}_{nm} + p_n h_{nn}^2 \tilde{\kappa}_{nn}. \quad (16)$$

Let us denote the numerator and denominator of the fraction in (13) as $\mathcal{N}_{nm}(\mathbf{s})$ and $\mathcal{D}_{nm}(\mathbf{s})$ respectively. Thus, $\mathcal{N}_{nm}(\mathbf{s})$ can be written in a simplified form as

$$\mathcal{N}_{nm}(\mathbf{s}) = \alpha_{nm} - \mathbf{s}^T \mathbf{U}_{nm} \mathbf{s} - \mathbf{s}^T \mathbf{c}_{nm} - |v_{nm}|^2, \quad (17)$$

where \mathbf{U}_{nm} is the outer product of \mathbf{u}_{nm} , i.e., $\mathbf{U}_{nm} = \mathbf{u}_{nm} \mathbf{u}_{nm}^H$ and $\mathbf{c}_{nm} = v_{nm}^* \mathbf{u}_{nm} + v_{nm} \mathbf{u}_{nm}^*$. The positive semi-definite characteristics of \mathbf{U}_{nm} render the concavity characteristic of \mathcal{N}_{nm} in \mathbf{s} . Similarly, $\mathcal{D}_{nm}(\mathbf{s})$ can be rewritten as

$$\mathcal{D}_{nm}(\mathbf{s}) = \beta_{nm} - \mathbf{s}^T \mathbf{W}_{nm} \mathbf{s} - \mathbf{s}^T \mathbf{d}_{nm} - |v_{nm}|^2, \quad (18)$$

with $\mathbf{d}_{nm} = v_{nm}^* \mathbf{w}_{nm} + v_{nm} \mathbf{w}_{nm}^*$ and $\mathbf{W}_{nm} = \mathbf{w}_{nm} \mathbf{w}_{nm}^H$, which implies the concavity of \mathcal{D}_{nm} in \mathbf{s} thanks to the positive semi-definite properties of \mathbf{W}_{nm} . Therefore, using the aforementioned representations, optimization problem **P1** can be equivalently written as

$$\begin{aligned} \mathbf{P2} : \quad & \text{maximize} \quad \min_n \left\{ \frac{B}{2} \log_2 \frac{\mathcal{N}_{nm}(\mathbf{s})}{\mathcal{D}_{nm}(\mathbf{s})} \right\} \\ & \text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \quad \forall m, \end{aligned} \quad (19)$$

where $\mathbf{P}_m = \text{diag} \left\{ \langle \langle [1 \ 1]^T \rangle_{(Az, 2k)} \rangle_{(Rs, 2m)} \right\}$.

The max-min fractional problem in **P2** can be efficiently solved by exploiting the properties of the logarithmic function. In addition, we can transform the secondary minimization problem into the following maximization problem

$$\begin{aligned} \mathbf{P3} : \quad & \text{maximize} \quad - \max_n \left\{ -\frac{B}{2} \log_2 \mathcal{N}_{nm}(\mathbf{s}) + \frac{B}{2} \log_2 \mathcal{D}_{nm}(\mathbf{s}) \right\} \\ & \text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \quad \forall m. \end{aligned} \quad (20)$$

Evidently, $\log_2 \mathcal{N}_{nm}(\mathbf{s})$ and $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ are concave functions owing to the the positive-concave nature of $\mathcal{N}_{nm}(\mathbf{s})$ and $\mathcal{D}_{nm}(\mathbf{s})$, respectively. Therefore, the subtractive form of the objective function in **P3** is universally known as DC-programming (difference of concave) and cannot be handled straightforwardly. Thus, we employ sequential convex programming (SCP) to efficiently transform **P3** into a iterative convex problems which can be optimally solved in each iteration [61]. In this approach, we use the affine Taylor series approximation of the function $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ to yield a convex objective function. The first-order Taylor series expansion of the function $g(x)$ at point $x^{(k)}$ is given by

$$\hat{g}(x, x^{(k)}) = g(x^{(k)}) + \nabla g(x^{(k)})^T (x - x^{(k)}). \quad (21)$$

Thus, by employing the same expansion for the concave function $\log_2 \mathcal{D}_{nm}(\mathbf{s})$ gives an affine approximation at $\mathbf{s}^{(i)}$ as

$$\tilde{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) = \frac{B}{2 \ln 2} \left(\ln \mathcal{D}_{nm}(\mathbf{s}^{(i)}) + \frac{\nabla^T \mathcal{D}_{nm}(\mathbf{s}^{(i)})}{\mathcal{D}_{nm}(\mathbf{s}^{(i)})} (\mathbf{s} - \mathbf{s}^{(i)}) \right), \quad (22)$$

where $\nabla \mathcal{D}_{nm}(\mathbf{s}^{(i)}) = -(\mathbf{W}_{nm} + \mathbf{W}_{nm}^T) \mathbf{s}^{(i)} - \mathbf{d}_{nm}$ is the gradient of $\mathcal{D}_{nm}(\mathbf{s})$ evaluated at $\mathbf{s}^{(i)}$. It is important to note that no trust region is required as $\tilde{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) \leq \mathcal{D}_{nm}(\mathbf{s}^{(i)})$ [61]. Thus, **P3**

can be convexified using the aforementioned procedure giving the following problem that needs to be solved successively while updating $\mathbf{s}^{(i)}$,

$$\begin{aligned} \mathbf{P4}: \quad & \text{maximize} \quad -\max_n \left\{ -\frac{B}{2} \log_2 \mathcal{N}_{nm}(\mathbf{s}) + \check{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) \right\} \\ & \text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \forall m. \end{aligned} \quad (23)$$

Given, the convex objective function and convexity preservation by the point-wise maximization, the primary maximization problem can be equivalently written as follows [62]

$$\begin{aligned} \mathbf{P5}: \quad & \text{minimize} \quad \max_n \left\{ -\frac{B}{2} \log_2 \mathcal{N}_{nm}(\mathbf{s}) + \check{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) \right\} \\ & \text{subject to} \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \forall m. \end{aligned} \quad (24)$$

Algorithm 1 Sequential Convex Programming

- 1: **Initialize** $i \leftarrow 0$, $\epsilon \leftarrow \infty$ and **Set** tolerance δ
 - 2: **Choose** feasible starting point $\mathbf{s}^{(i)}$
 - 3: **while** $\epsilon \geq \delta$ **do**
 - 4: Evaluate $\check{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) \forall m \setminus \{k\}$
 - 5: Solve **P6** and obtain \mathbf{s} using $\mathbf{s}^{(i)}$
 - 6: $\mathbf{s}^{(i+1)} \leftarrow \mathbf{s}$
 - 7: Update $\epsilon \leftarrow |\mathbf{s}^{i+1} - \mathbf{s}^i|$
 - 8: $i \leftarrow i + 1$
 - 9: **end while**
 - 10: $\mathbf{s}^* \leftarrow \mathbf{s}^{i+1}$
 - 11: $R_T = \min_n \{R_{nm}(\mathbf{s}^*)\}$
-

The formulated problem **P5** yields an optimal solution for a given $\mathbf{s}^{(i)}$ as it is the minimization of a convex function pertaining to the convexity preservation by point-wise maximization. One way to solve **P5** is by introducing an auxiliary variable τ in order to capture the point-wise maximization problem as follows

$$\begin{aligned}
\mathbf{P6} : \quad & \text{minimize} \quad \tau \\
& \text{subject to} \quad -\frac{B}{2}\log_2\mathcal{N}_{nm}(\mathbf{s}) + \check{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)}) \leq \tau, \quad \forall m \setminus \{k\} \\
& \quad \quad \quad -\frac{B}{2}\log_2\mathcal{N}_{nm}(\mathbf{s}) + \frac{B}{2}\log_2\mathcal{D}_{nm}(\mathbf{s}) \leq \tau, \quad m = k \\
& \quad \quad \quad 0 \leq \mathbf{s}^T \mathbf{P}_m \mathbf{s} \leq \sigma_m^2, \quad , \forall m. \tag{25}
\end{aligned}$$

Therefore, the solution of the proposed optimization problem reduces to solving **P6** iteratively using the SCP as discussed in the following subsection.

1) *Centralized Joint Algorithm:* To solve **P1** or equivalently **P2**, we use the SCP that deals with a convexified version of the difference of convex/concave problem **P3** in each iteration developing Algorithm 1. The proposed algorithm starts with a feasible starting point $\mathbf{s}^{(i)}$ to find the affine approximation of $\frac{1}{2}\log_2\mathcal{D}_{nm}(\mathbf{s})$ for all transmitters, i.e., R_m and $m < k$. The last node does not suffer from self-interference, which renders constant $\frac{1}{2}\log_2\mathcal{D}_{nm}(\mathbf{s})$ at $m = k$. Then, the affine approximation $\check{\mathcal{D}}_{nm}(\mathbf{s}, \mathbf{s}^{(i)})$ is used to optimally solve a quadratic-constraint linear-programming (QCLP) problem defined in **P6** using any available convex optimization solvers such as CVX-MATLAB employing interior-point method. Next, the solution of the QCLP problem is used to update \mathbf{s} obtaining $\mathbf{s}^{(i+1)}$, which is the starting point for next iteration. The algorithm solves successive convex QCLPs and updates the solution values in each iteration until the desired stopping condition is met. The stopping convergence criterion is when the absolute difference between two successive solutions is less than a predefined threshold δ .

The obtained solution vector \mathbf{s} contains the real and imaginary components of the transmit pseudo-variances for all transmitting nodes. Thus, the maximized end-to-end achievable rate R_T can be computed using (6) and (13). Despite of the efficacy of the joint-optimization, it requires a centralized network realization, where all nodes share their information with a central node prior to the transmission. Then, the signal design is carried out at the central node in order to update the relaying nodes with the optimized signal parameters.

B. Distributed Optimization

The implementation of the joint optimization requires a centralized network realization. However, the centralized realization may not be suitable in different scenarios such as:

- **Communication Overhead:** A network involving large number of relays to establish a communication link between two nodes, the pilot CSI information from all intermediate nodes is sent to a central node on the same communication links in an accumulative fashion. Then the central node, with high performance capabilities, performs transmit optimization and distributes the optimized parameter values back to each node. This leads to communication overhead and it increases with increasing number of relays which requires a powerful central node to accommodate the data processing and computation complexity burden. However, the communication overhead decreases in a distributed mechanism as the entire network is divided into smaller clusters with fewer number of relays and each cluster locally performs the transmit optimization independent of other clusters.
- **Time Delays:** Intuitively, we employ relay to establish a communication link between any two distant nodes in the absence of a direct link. This same communication link is utilized to transmit CSI from each node to the central node using intermediate nodes. Evidently, more time slots are required to transmit the pilot CSI from a large number of nodes to a central node which not only results in large delays but may also cause error propagation. Then the central node processes the received information, performs joint optimization and distributes the optimized parameter values back to each node through the intermediate nodes. Moreover, the processing/computational delay at the central node also increases with the increasing nodes. Joint optimization of the transmit parameters of all involved nodes increases the dimension of the optimization problem and hence adds to computational delays relative to the distributed algorithm within one cluster with fewer number of nodes. For large networks, this round trip time and computational delay may exceed channel coherence time. This will render sub-optimal transmission in the new channel state as the parameters were optimized for the previous CSI.

Therefore, to address the practical limitations of the centralized realization, reduce the communication/processing overhead on one central node and avoid the excessive time delays, the distributed optimization framework is proposed.

1) *Distributed Algorithm:* In the distributed framework, we aim to group the neighboring nodes into clusters of equal sizes, having N nodes and $N - 1$ hops, in order to equally distribute the computational load among all clusters. Each cluster acts as a whole system and locally runs the joint-optimization algorithm to maximize the end-to-end achievable rate within that cluster. However, the scenarios arise when the total number of nodes is not an integer multiple of the

desired cluster size and thus they cannot be equally grouped into the clusters of same sizes. Consider a system with k -relays between the source and the destination. We propose to divide this system into C equal sized clusters, where each cluster has N nodes. C is found to be expressed as

$$C = \left\lfloor \frac{k+1}{N-1} \right\rfloor. \quad (26)$$

which leads to N_s unassigned nodes where N_s is smaller than N . In the following, we propose two grouping configurations to deal with the N_s nodes:

- Group the N_s nodes to form an additional small-cluster, thus the system will have a total of $C + 1$ clusters. As such, N_s nodes are found from:

$$N_s = k - \left\lfloor \frac{k+1}{N-1} \right\rfloor (N-1) + 2. \quad (27)$$

- Group the remaining N_s nodes with a pre-defined cluster. Thus, the number of nodes in the new re-defined cluster becomes N_b and is found to be

$$N_b = N + N_s - 1. \quad (28)$$

The performance of the distributed realization approach can be improved by smartly choosing the cluster size and their placement in the network. A heuristic approach is to formulate this odd sized cluster either at the transmitter side involving source node or at the receiver end involving destination node. However, it is important to note that the cluster formulation yields sub-optimal solution owing to the neglected RSI at the end nodes. Therefore, the big odd cluster is expected to outperform the small odd cluster and even the regular cluster pertaining to the inclusion of RSI at all intermediate nodes. Thus, a preferred approach is to place the small cluster at the strongest channel location so that its compromised transmission parameters are not detrimental for the overall end-to-end achievable rate or formulate the big odd cluster at the weakest channel location as it will contain the bottleneck link dictating end-to-end achievable rate. This concept has been supported and validated using the simulation results as discussed in Section V.

For comprehensive illustration, consider the distributed network scheme presented in Fig. 2. We aim to divide this system having $k = 4$ relays into $N = 3$ sized clusters, which is clearly not possible. Therefore, one approach is to group the remaining $N_s = 2$ nodes in another cluster which can either be placed in the middle or at the destination as shown in Fig. 2(left). Another possibility is to place this at the transmitting end. As discussed, we prefer to formulate this cluster at the $\max(|h_{10}|, |h_{32}|, |h_{54}|)$. So that the compromised transmission parameters due to

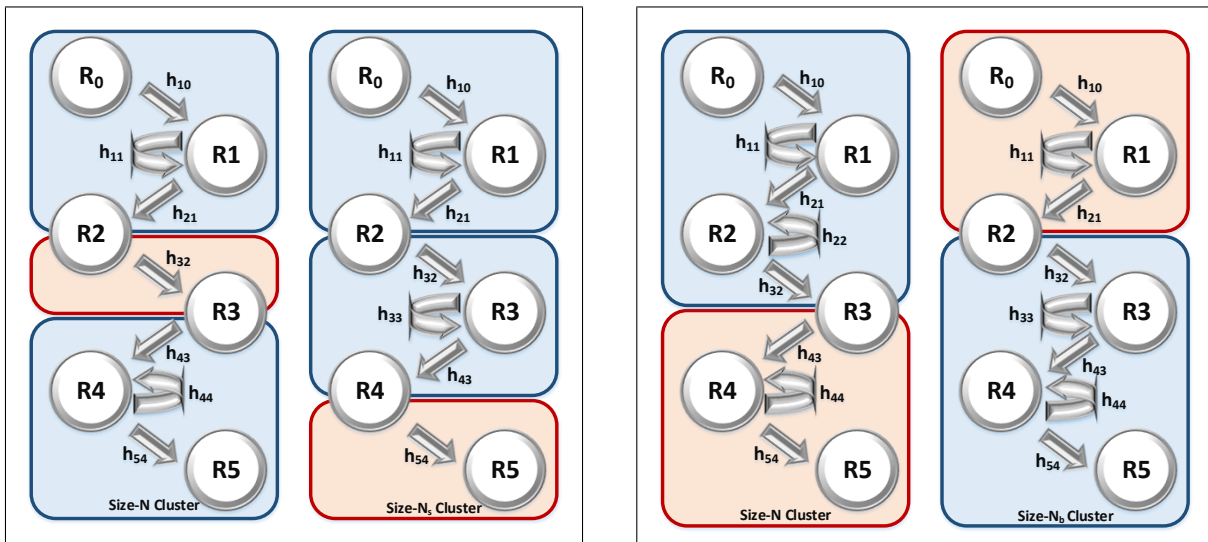


Fig. 2. $C + 1$ clusters with one N_s cluster and its two possible placements (Left) . C clusters with one N_b cluster and its two possible placements (Right).

neglected RSI do not dominate the system performance, which is dictated by the minimum rate link. Another approach is to group the remaining $N_s = 2$ nodes with another cluster formulating a $N_b = 4$ sized-cluster which can either be placed at the start or at the end as shown in Fig. 2(right). Again, the preferred approach is to place this big cluster around transmitter if the $\min(|h_{10}|, |h_{21}|) \leq \min(|h_{43}|, |h_{54}|)$, else place it at the end to accommodate destination node. Thus, the bigger cluster will host the bottleneck link and the conforming transmission parameters, owing to the inclusion of RSI at larger number of intermediate nodes, will dictate the overall system throughput.

It is important to highlight that the above mentioned approach requires a rough idea of the strong and/or weak channel gains throughout the span of the system to practically locate the odd cluster. The small cluster and big cluster scenarios are separately dealt in Algorithm 2 and 3 respectively. As mentioned earlier, the small cluster size leads to inefficient solution therefore we place this small cluster at a location which offers relatively higher channel gain. This is advantageous as the cluster with better channel conditions doesn't dictate the overall system achievable rate. Similarly, we place big cluster at a weaker channel location for more accurate solution parameters as it will dictate the overall system rate. The following two algorithms are enumerated to practically implement the distributed networking approach.

In Algorithm 2, we begin by choosing a cluster size N and divide the entire system into C

Algorithm 2 Distributed Algorithm with $C + 1$ clusters

- 1: **Choose** Cluster size N
 - 2: **Compute** Smaller cluster size N_s using (27).
 - 3: **Formulate** the small cluster at the strongest channel location.
 - 4: **Initialize** Cluster counter $j \leftarrow 1$
 - 5: **while** $j \leq C + 1$ **do**
 - 6: **if** Before N_s **then**
 - 7: Solve **P2** for $m = (j - 1)(N - 1), \dots, (j)(N - 1) - 1$.
 - 8: **end if**
 - 9: **if** At N_s **then**
 - 10: Solve **P2** for $m = (j - 1)(N - 1), \dots, (j)(N - 1) - (N - N_s) - 1$.
 - 11: **end if**
 - 12: **if** Beyond N_s **then**
 - 13: Solve **P2** for $m = (j - 1)(N - 1) - (N - N_s), \dots, (j - 1)(N - 1) + N_s - 2$.
 - 14: **end if**
 - 15: **Distribute** the locally optimized transmit parameters within cluster.
 - 16: $j \leftarrow j + 1$
 - 17: **end while**
-

number of clusters each having N nodes. The remaining nodes are grouped into a small cluster of size N_s and are chosen to have relatively high channel gain. Then, Algorithm 2 groups the clusters into three main groups with a total of $C + 1$ clusters. The first group consists of all clusters before the N_s cluster. The second group comprises only the N_s cluster. Finally, the last group comprises of all the clusters after the N_s cluster. The physical intuition behind the three mentioned groups in the algorithm is to provide the appropriate indexing of the involved transmitting nodes in each cluster for their transmit parameter optimization. The transmit parameters are locally designed in an individual cluster by solving optimization problem **P2** for the corresponding range of m -transmitters. It is worth mentioning that the problem **P2** is solved by iteratively solving the problem **P6** using Algorithm 1. The optimized variables are then distributed to each node within the cluster and the process continues.

Algorithm 3 realizes the C cluster scenario, where the N_s cluster is merged with a regular

Algorithm 3 Distributed Algorithm with C clusters

- 1: **Choose** Cluster size N
 - 2: **Compute** bigger cluster size N_b using (28).
 - 3: **Formulate** the big cluster at the weakest channel location.
 - 4: **Initialize** Cluster counter $j \leftarrow 1$
 - 5: **while** $j \leq C$ **do**
 - 6: **if** Before N_b **then**
 - 7: Solve **P2** for $m = (j - 1)(N - 1), \dots, (j)(N - 1) - 1$.
 - 8: **end if**
 - 9: **if** At N_b **then**
 - 10: Solve **P2** for $m = (j - 1)(N - 1), \dots, (j - 1)(N - 1) + N_b - 2$.
 - 11: **end if**
 - 12: **if** Beyond N_b **then**
 - 13: Solve **P2** for $m = (j - 1)(N - 1) + N_b - N, \dots, (j - 1)(N - 1) + N_b - 2$.
 - 14: **end if**
 - 15: **Distribute** the locally optimized transmit parameters within cluster.
 - 16: $j \leftarrow j + 1$
 - 17: **end while**
-

cluster to formulate N_b sized cluster. Similar to Algorithm 2, Algorithm 3 deals with three groups of clusters by scanning the C participating clusters. In the absence of CSI, the remaining cluster can be randomly placed anywhere in the network. Although the division into clusters reduces the round trip delays, communication overhead, and computational complexity, but it does so at the expense of deviated solution parameters. The distributed solution ignores the drastic effects of RSI at the destination node in each cluster. Thus, the solution of the distributed setup is expected to deviate from the joint optimization one depending on the cluster size. The bigger the cluster size, the closer the distributed optimization solution to the joint optimization one, as it accommodates RSI of the intermediate nodes and vice versa. Moreover, the distributed optimization focuses on the minimum rate performance of local cluster without considering other links, which may causing deviation from the joint optimization solution. This portrays a trade off between the distributed computational load and the optimal performance.

TABLE I
COMPLEXITY ANALYSIS OF THE PROPOSED ALGORITHMS

Complexity Analysis			
Joint Algorithm	Distributed Algorithm		
$O(\alpha I_1^{\text{SCP}} k^3)$	Regular N -sized Cluster	Small N_s -sized Cluster	Big N_b -sized Cluster
	$O(\alpha I_2^{\text{SCP}} N^3)$	$O(\alpha I_{2a}^{\text{SCP}} N_s^3)$	$O(\alpha I_{2b}^{\text{SCP}} N_b^3)$

C. Complexity Analysis

The computational complexity analysis of the proposed algorithms has been carried out in this subsection. The proposed joint and distributed transmit optimization algorithms depict a trade-off between performance and the computational complexity assuming negligible round-trip delays and the high end computational capabilities of the central node. The computational complexity of these strategies can be expressed as follows:

- Joint Approach: $O(I_1^{\text{SCP}} (N_1^2 k + \alpha \max(N_1^3, N_1^2 M_1, F_1)))$
- Distributed Approach: $O(I_2^{\text{SCP}} (N_2^2 (\Gamma - 2) + \alpha \max(N_2^3, N_2^2 M_2, F_2)))$

where I_x^{SCP} is the number of SCP iterations before convergence ($x = 1$ for joint algorithm, $x = 2$ for distributed algorithm with regular clusters of size N , $x = 2a$ for distributed algorithm with small cluster of size N_s and $x = 2b$ for distributed algorithm with big cluster of size N_b). I_x^{SCP} iterations are the increasing function of the cluster-size as it adds more dimensions to the optimization problem. In addition, the total number of optimization variables in joint and distributed optimization algorithms are given as $N_1 = 2k + 2$ and $N_2 = 2\Gamma - 2$, respectively. Where k is the total number of relays and Γ is the number of nodes in a given distributed cluster. As for α , it is assumed to be between 10 and 100 for the interior point method [62]. Moreover M_1 and M_2 are the number of inequalities representing constraints of joint and distributed optimization problems, respectively, defined as $M_1 = 2(k + 1)$ and $M_2 = 2(\Gamma - 1)$, respectively. Furthermore, $F_1 = N_1(1 + M_1) + N_1^2 M_1$ and $F_2 = N_2(1 + M_2) + N_2^2 M_2$ are the costs of evaluating the first and second derivatives of the objective and constraint functions in joint and distributed algorithm, respectively. The computational complexities of joint and distributed algorithm are further simplified in Appendix A and Appendix B, respectively. The simplified complexity analysis is presented in Table I for joint-algorithm with k relays and distributed-algorithm with $\Gamma = N$ sized regular cluster, $\Gamma = N_s$ sized small cluster, and $\Gamma = N_b$ sized big

cluster.

Evidently, the trend $1 \leq N_s \leq N \leq N_b \leq k$ follows from (27) and (28), depicts the least complexity for the N_s -cluster distributed algorithm relative to the N -cluster distributed algorithm. Moreover, N_b -cluster distributed algorithm depicts the most complexity of all distributed algorithms. However, the relation $\Gamma = (k+C+1)/C$ demonstrates that $\Gamma \ll k$. Thus, validating that the computational complexity of the distributed algorithm in a cluster of size Γ is far less than that of the joint optimization algorithm involving k relays.

V. NUMERICAL AND SIMULATION RESULTS

In this section, we quantify the gain reaped by employing IGS transmission scheme in place of PGS scheme for the adopted MH-DF-FDR system suffering from RSI and HWD. Besides studying the degradation effect caused by these interferences, we also investigate the impact of multiple relays and the cluster size in distributed optimization approach on the overall system performance. Moreover, we compare the performance of the two proposed algorithm for the distributed approach with N_b and N_s cluster sizes and their respective placement in the system network.

As for the simulation parameters, we assume unity transmit power from all transmitters for simplicity, noise variance $\sigma_z^2 = 1$ and HWD level $\kappa_{nm} = 1$ along with the impropriety level $|\tilde{\kappa}_{nm}| = 0.9$ at all participating nodes.

Moreover, h_{nm} is modeled as a slowly-varying Rayleigh flat-fading channel of the m - n link with $h_{nm} \sim \mathcal{CN}(0, \pi_{nm}, 0)$ ³. Furthermore, we assume 25dB SNR for m - n link and the RSI of 10dB unless otherwise specified.

A. Effect of RSI and HWD

At first we study the performance degradation caused by the RSI and HWD in Fig. 3 and Fig. 4 respectively. In the first simulation example, we study the advantage of employing IGS in suppressing the RSI effect on the average achievable end-to-end rate for a dual-hop FDR system as shown in Fig. 3. Average rate is observed at various RSI gains π_{rr} ranging from 0dB to 25dB for three different HWD levels. We assume equal HWD levels at source, relay

³ It is important to highlight that the presented technical contribution holds true for any form of fading including Rayleigh, Ricean or Nakagami. It is because of the fact that the derived framework does not depend on the statistical characteristics of the given channel model.

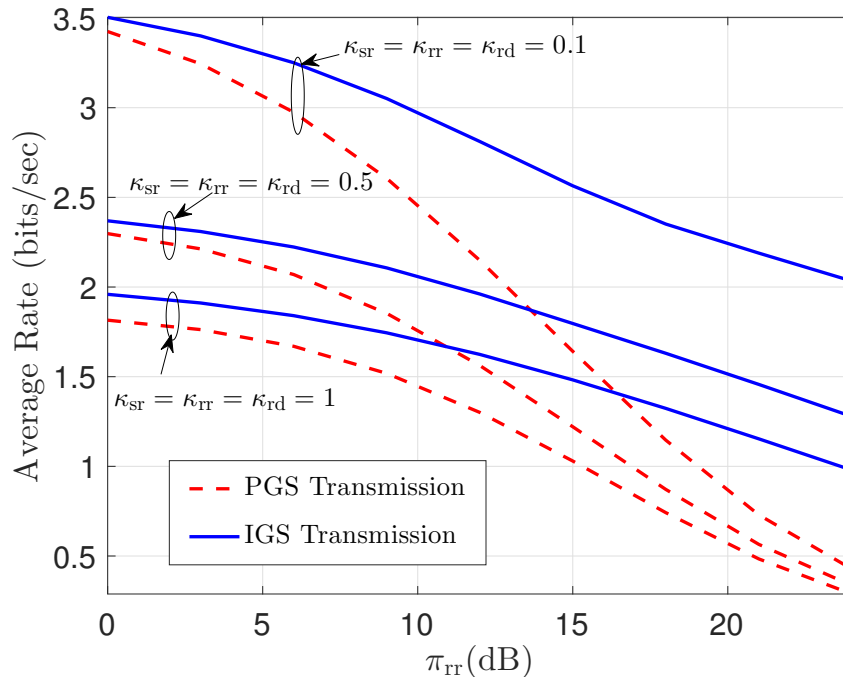


Fig. 3. Average rate vs. RSI π_{IR} for PGS and IGS schemes assuming different distortion levels.

and destination as presented by $\kappa_{\text{sr}} = \kappa_{\text{rr}} = \kappa_{\text{rd}}$ for simplicity. Evidently, the increasing self-interference severely degrades the achievable rate performance. In addition, increasing HWD κ_{nm} from 0.1 to 1 also deteriorates the system performance. Interestingly, the proposed IGS scheme is capable of providing significant performance enhancement at lower-residual HWD levels, assuming effective joint compensation of HWD and SI, for the entire range of RSI levels as shown in Fig. 3. Similar results have been demonstrated in [35] to emphasize the significance of employing IGS transmission in an attempt to alleviate the RSI adverse effect in FDR system considering ideal transceivers. It is important to highlight that the PGS scheme undergoes saturation at higher RSI levels irrespective of the HWD level. Thus it can be safely concluded that the RSI dominates in degrading the rate performance for PGS. On the other hand, IGS scheme efficiently mitigates the RSI impact and reduces the rate degradation. The best relative improvement is achieved at high RSI and low HWD levels.

Secondly, we study the degradation effect of HWD on the average achievable rate for various impropriety levels and the relative performance gains obtained by the proposed joint IGS scheme and a sub-optimal less-complex distributed IGS approach over the conventional PGS scheme in

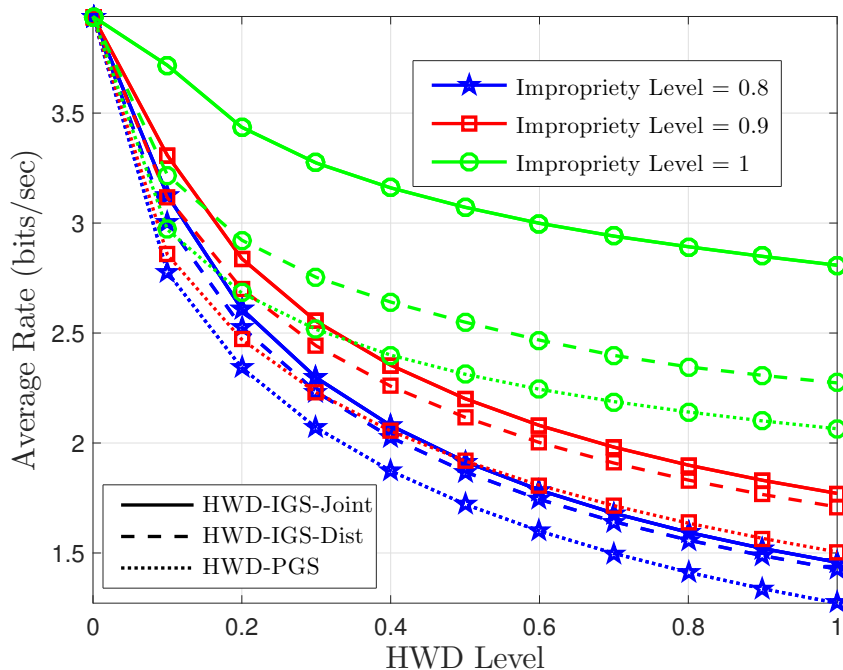


Fig. 4. Average achievable rate versus HWD level with equal impairment levels at all nodes κ_{mn}

Fig. 4. We assume a MH system incorporating 3 relays between the source and destination with favorable channel gains of $h_{nm}=30\text{dB}$ at each participating link. Clearly, the rate performance drastically deteriorates with increasing HWD variance, from 0 indicating the ideal hardware to 1 indicating the maximal impairment level⁴ even at very good channels gain. Moreover, the joint IGS optimally mitigates the distortion impact relative to the PGS scheme especially for higher improprity levels. In addition, the distributed approach with cluster size $N = 2$ locally optimizes the transmission parameters reducing complexity and undesired communication overhead and delays. Distributed approach performs sub-optimal to joint IGS however it is still way better than the traditional PGS. Interestingly, the distributed approach performs close to optimal joint IGS for lower improprity levels.

B. Number of Participating Relays

Next, we investigate the effect of increasing relays on the end-to-end achievable rate in the

⁴Maximal impairment level is attained when the additive distortion power becomes equal to the transmitted signal power in the absence of any other mitigation strategy

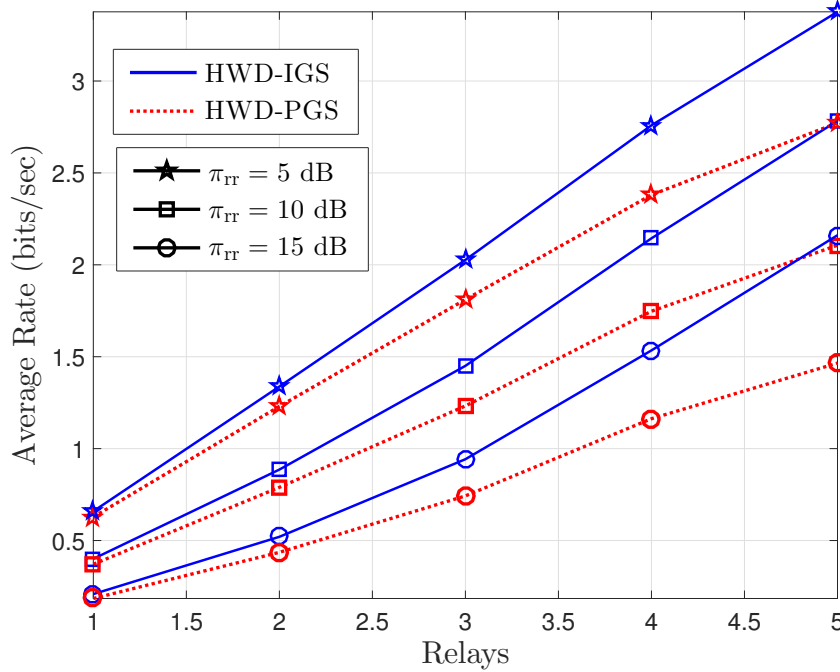


Fig. 5. Average achievable rate versus number of participating relays.

absence of a direct link between a given source and destination. We assume a limited power budget for all participating transmitters including source and relays. Thus, increasing the number of relays decreases the power transmitted by one transmitter yielding negligible inter-relay interference. Furthermore, we have adopted a Rayleigh fading and distance path-loss model for practical realization. This implies increasing channel strength between any two nodes with increasing relays. We assume increasing number of relays ranging from 1 to 5 with increasing channel strengths from 8dB SNR to 35dB SNR and decreasing transmit power from 3W to 1W per transmitter. We further assume three levels of RSI i.e., 5dB, 10dB and 15dB as shown in Fig. 5. Intuitively, the absence of inter-relay interference and better link strength with increasing number of relays guarantees an increase in the achievable rate. However, the presence of HWD and limited power budget limits this performance gain. Moreover, increasing RSI drastically degrades the average rate performance especially for PGS and higher number of participating relays. Conclusively, joint IGS scheme efficiently mitigates the interference effects as compared to the existing PGS with any given number of relays and RSI levels.

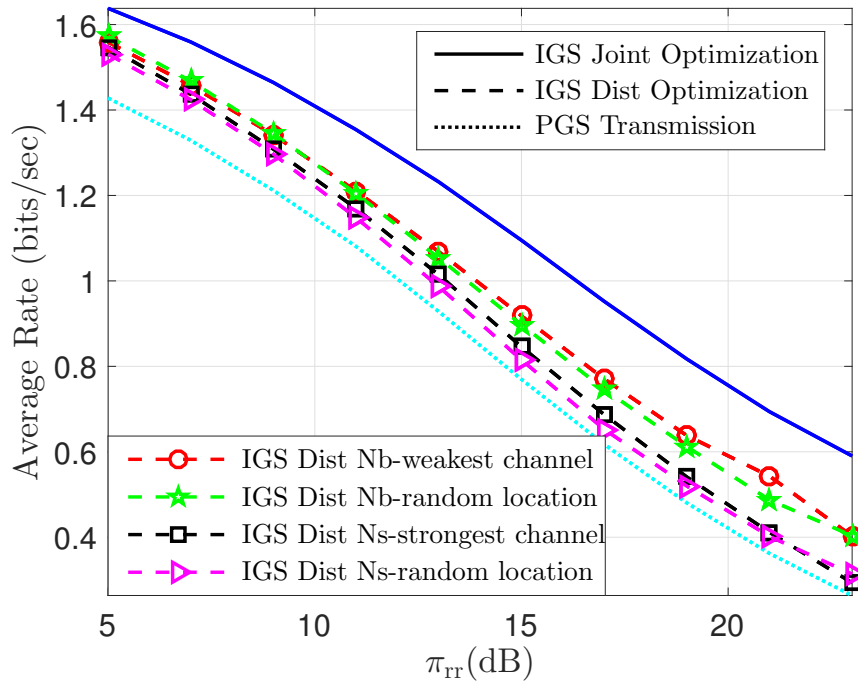


Fig. 6. Average achievable rate versus self-interference employing different distributive algorithms .

C. Impact of Cluster size and Placement in Distributed Algorithms:

Another simulation example in Fig. 6 illustrates the system achievable rate attained by various forms of distributive algorithms keeping the joint-IGS scheme as a benchmark. A MH system accommodating 4-relays between the source and destination is optimized using four different forms of $N = 3$ distributive framework employing algorithm 2 and 3. IGS distributed algorithm with N_b cluster divides the 5-link network into a cluster of size $N = 3$ and groups the remaining nodes in a bigger cluster of size $N_b = 4$. Further division is based on the placement of this N_b cluster, it is evident from Fig. 6 that the odd cluster placement at weak channel location (IGS Dist Nb) outperforms the random placement in the absence of CSI. Similarly, IGS distributed algorithm with N_s cluster divides the 5-link network into 2 clusters of size $N = 3$ and the remaining link is isolated. Again the placement of the segregated link at the strongest channel location (IGS Dist Ns) performs better than the random placement at any other location. It is worth noting that the N_b cluster formation outperforms the N_s cluster formation irrespective of the cluster placement. Therefore, N_b distributed algorithm is the preferred choice when feasible. Evidently, all the proposed IGS schemes provide significant rate compensation/improvement at

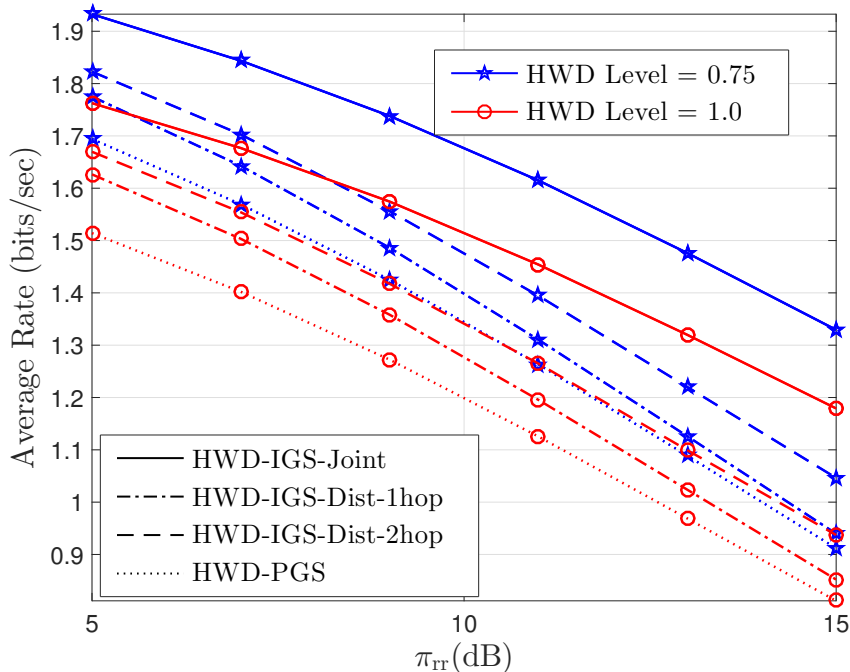


Fig. 7. Average achievable rate versus self-interference for various cluster sizes.

different interference levels as compared to its counterpart PGS scheme.

Finally, the impact of cluster size is analyzed on the end-to-end achievable rate for two different HWD levels in Fig. 7. 3-relay FDR system under HWD and RSI is locally optimized using 3-different cluster sizes $N = 2$ or 1-hop, $N = 3$ or 2-hop, and $N = 5$ or 4-hop in addition to the joint optimization scheme. It is apparent from Fig. 7 that the bigger the cluster size, the closer rate performance of the distributed to the joint optimization one. This proves the previously discussed intuition in Section IV-B. However, there is a trade-off between the cluster size and the system complexity, round-trip delays and communication overhead. Therefore, the cluster size can be decided according to a given system configuration.

Numerical and simulation results clearly advocate the benefits of deploying various forms of the IGS scheme over the PGS scheme to improve the system performance in terms of end-to-end achievable rate for various levels of RSI and HWD in the MH-FDR system. Moreover, the distributed approach performs closer to the joint optimization approach for big cluster sizes and lesser number of clusters as shown in Fig. 6 and Fig. 7, respectively. It is owing to the accommodation of the RSI effects at all intermediate nodes.

The theoretical limits attained by IGS transmission for the achievable rate of a MH-DF-FDR system suffering from HWDs establish the performance limits that can be achieved through the modern communication standards. The existing techniques can achieve the performance demonstrated by the PGS scheme. However, this work motivates the future research to propose appropriate adaptive coding and modulation which can achieve the IGS performance.

VI. CONCLUSION

In this paper, we analyzed the effectiveness of using IGS scheme in MH-DF-FDR systems under HWD and RSI. To this end, we expressed the achievable rate for the underlying system and tuned the IGS pseudo-variance to maximize the end-to-end achievable rate. We presented, analyzed and illustrated two realization schemes named as joint and distributed optimization schemes. Distributed-IGS is further categorized as per the cluster size as well as its relative position in the system network. Distinct forms of IGS-scheme can be adopted for suitable system configurations. For a small system configuration with fewer hops joint-IGS is the preferred choice. However, for a larger-hops system, the joint-IGS renders sub-optimal results pertaining to the inevitable processing and round-trip delays back and forth from the central-node at the cost of increased system complexity and communication overhead. Therefore, distributed-IGS is the preferred approach for large system configurations. Furthermore, distributed-IGS with bigger odd-cluster along with the optimal cluster placement is the preferred choice as per the acquired simulation results. In a nutshell, all forms of IGS are proven to be the promising candidates for next generation networks that can significantly improve the overall achievable rate under various HWD and RSI levels, which have asymmetric signatures on the useful signal.

VII. ACKNOWLEDGMENTS

This work was funded under grant #AT-35-59 from King Abdulaziz City of Science and Technology.

APPENDIX A

COMPUTATIONAL COMPLEXITY OF THE JOINT ALGORITHM

Using the definitions in Section IV-C, the computational complexity of the joint algorithm can be simplified as follows:

$$O(I_1^{\text{SCP}}(N_1^2 k + \alpha \max(N_1^3, N_1^2 M_1, F_1))) \quad (29)$$

Considering $N_1^2 k \leq \alpha \max(N_1^3, N_1^2 M_1, F_1)$, we can simplify the complexity expression as

$$O(\alpha I_1^{\text{SCP}} \max(N_1^3, N_1^2 M_1, F_1)) \quad (30)$$

Using $N_1^3 = N_1^2 M_1 = 8k^3 + 24k^2 + 24k + 8$ and $F_1 = 8k^3 + 28k^2 + 34k + 14$, depicts $F_1 = \max(N_1^3, N_1^2 M_1, F_1)$. Thus, the complexity reduces to $O(\alpha I_1^{\text{SCP}} F_1)$. Considering the dominant term in F_1 renders the desired complexity of order $O(\alpha I_1^{\text{SCP}} k^3)$.

APPENDIX B

COMPUTATIONAL COMPLEXITY OF THE DISTRIBUTED ALGORITHM

Similarly, the computational complexity of the distributed algorithm with cluster size Γ can be simplified to $O(\alpha I_2^{\text{SCP}} \max(N_2^3, N_2^2 M_2, F_2))$. Following the similar steps as in Appendix A with $F_2 = \max(N_2^3, N_2^2 M_2, F_2)$ and using $F_2 = 8\Gamma^3 - 20\Gamma^2 + 18\Gamma - 6$, yields the following complexity of order

$$O(\alpha I_2^{\text{SCP}} F_2) = O(\alpha I_2^{\text{SCP}} \Gamma^3) \quad (31)$$

REFERENCES

- [1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.
- [2] R. Q. Hu and Y. Qian, "An energy efficient and spectrum efficient wireless heterogeneous network framework for 5G systems," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 94–101, May 2014.
- [3] V. Genc, S. Murphy, Y. Yu, and J. Murphy, "IEEE 802.16 J relay-based wireless access networks: an Overview," *IEEE Wireless Commun. Mag.*, vol. 15, no. 5, Oct. 2008.
- [4] O. Amin, S. Bavarian, and L. Lampe, "Cooperative techniques for energy-efficient wireless communications," in *Green Radio Communication Networks*. Cambridge University Press, 2012, pp. 125–149.
- [5] G. Farhadi and N. C. Beaulieu, "On the ergodic capacity of multi-hop wireless relaying systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, Apr. 2009.
- [6] H. Ju, S. Lim, D. Kim, H. V. Poor, and D. Hong, "Full duplexity in beamforming-based multi-hop relay networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 8, pp. 1554–1565, Sep. 2012.
- [7] L. Kong, L. Ye, F. Wu, M. Tao, G. Chen, and A. V. Vasilakos, "Autonomous relay for millimeter-wave wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 2127–2136, June 2017.
- [8] A. Masmoudi and T. Le-Ngoc, "Residual self-interference after cancellation in full-duplex systems," in *Communications (ICC), 2014 IEEE International Conference on*. Sydney, Australia: IEEE, June 2014, pp. 4680–4685.
- [9] 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, Nov. 2012.
- [10] E. Aryafar, M. A. Khojastepour, K. Sundaresan, S. Rangarajan, and M. Chiang, "MIDU: Enabling MIMO full duplex," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. Istanbul, Turkey: ACM, Aug. 2012, pp. 257–268.

- [11] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.
- [12] S. Li and R. D. Murch, "An investigation into baseband techniques for single-channel full-duplex wireless communication systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4794–4806, Sep. 2014.
- [13] D. Korpi, L. Anttila, V. Syrjälä, and M. Valkama, "Widely linear digital self-interference cancellation in direct-conversion full-duplex transceiver," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1674–1687, Sep. 2014.
- [14] T. T. Duy, T. Q. Duong, D. B. da Costa, V. N. Q. Bao, and M. Elkashlan, "Proactive relay selection with joint impact of hardware impairment and co-channel interference," *IEEE Trans. Commun.*, vol. 63, no. 5, pp. 1594–1606, May 2015.
- [15] J. Li, M. Matthaiou, and T. Svensson, "I/Q imbalance in AF dual-hop relaying: Performance analysis in Nakagami-m fading," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 836–847, Mar. 2014.
- [16] M. Mokhtar, A.-A. A. Boulogeorgos, G. K. Karagiannidis, and N. Al-Dhahir, "OFDM opportunistic relaying under joint transmit/receive I/Q imbalance," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1458–1468, May 2014.
- [17] M. Mokhtar, N. Al-Dhahir, and R. Hamila, "OFDM full-duplex DF relaying under I/Q imbalance and loopback self-interference," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6737–6741, Aug. 2016.
- [18] M. Mokhtar, A. Goma, and N. Al-Dhahir, "OFDM AF relaying under I/Q imbalance: Performance analysis and baseband compensation," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1304–1313, Apr. 2013.
- [19] X. Xia, D. Zhang, K. Xu, W. Ma, and Y. Xu, "Hardware impairments aware transceiver for full-duplex massive MIMO relaying," *IEEE Trans. Signal Process.*, vol. 63, no. 24, pp. 6565–6580, Dec. 2015.
- [20] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2017–2046, Feb. 2015.
- [21] I. Harjula, R. Wichman, K. Pajukoski, E. Lahetkangas, E. Tirola, and O. Tirkkonen, "Full duplex relaying for local area," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*. London, UK: IEEE, Sep. 2013, pp. 2684–2688.
- [22] D. Ramirez and B. Aazhang, "Optimal routing and power allocation for wireless networks with imperfect full-duplex nodes," *IEEE Trans. Wireless Commun.*, vol. 12, no. 9, pp. 4692–4704, Nov. 2013.
- [23] B. Mahboobi and M. Ardebilipour, "Joint power allocation and routing in full-duplex relay network: An outage probability approach," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1497–1500, July 2013.
- [24] T. K. Baranwal, D. S. Michalopoulos, and R. Schober, "Outage analysis of multihop full duplex relaying," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 63–66, Nov. 2013.
- [25] G. J. Gonzalez, F. H. Gregorio, J. Cousseau, T. Riihonen, and R. Wichman, "Performance analysis of full-duplex AF relaying with transceiver hardware impairments," in *European Wireless 2016; 22th European Wireless Conference; Proceedings of*. Oulu, Finland: VDE, May 2016, pp. 1–5.
- [26] O. Taghizadeh, A. C. Cirik, and R. Mathar, "Hardware impairments aware transceiver design for full-duplex amplify-and-forward MIMO relaying," *arXiv preprint arXiv:1703.10209*, 2017.
- [27] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Impact of improper Gaussian signaling on hardware impaired systems," in *Proc. 2017 IEEE Intern. Conf. Commun. (ICC 2017)*, Paris, France, May 2017.
- [28] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data: the theory of improper and noncircular signals*. Cambridge University Press, 2010.
- [29] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, June 2013.
- [30] S. Javed, O. Amin, S. S. Ikki, and M. S. Alouini, "On the achievable rate of hardware-impaired transceiver systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Singapore, Dec. 2017, pp. 1–6.

- [31] S. Javed, O. Amin, S. S. Ikki, and M.-S. Alouini, "Asymmetric hardware distortions in receive diversity systems: Outage performance analysis," *IEEE Access*, vol. 5, pp. 4492–4504, Feb. 2017.
- [32] O. Amin, W. Abediseid, and M.-S. Alouini, "Underlay cognitive radio systems with improper Gaussian signaling: Outage performance analysis," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, July 2016.
- [33] —, "Overlay spectrum sharing using improper Gaussian signaling," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 1, pp. 50–62, Jan. 2017.
- [34] M. Gaafar, M. G. Khafagy, O. Amin, and M.-S. Alouini, "Improper Gaussian signaling in Full-Duplex Relay channels with residual self-interference," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kuala Lumpur, May 2016.
- [35] M. Gaafar, M. G. Khafagy, O. Amin, R. F. Schaefer, and M.-S. Alouini, "Full-duplex relaying with improper Gaussian signaling over Nakagami- m fading channels," *IEEE Trans. Commun.*, vol. 66, no. 1, pp. 64–78, Jan. 2018.
- [36] M. Gaafar, O. Amin, A. Ikhlef, A. Chaaban, and M.-S. Alouini, "On alternate relaying with improper Gaussian signaling," *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1683–1686, Aug. 2016.
- [37] S. Javed, O. Amin, and M.-S. Alouini, "Full-duplex relaying under I/Q imbalance using improper Gaussian signaling," in *Proc. 42th Intern. Conf. Acoustics, Speech, and Signal Process. (ICASSP 2017)*, New Orleans, USA, Mar. 2017.
- [38] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.
- [39] E. Ollila, "On the circularity of a complex random variable," *IEEE Signal Process. Lett.*, vol. 15, pp. 841–844, Nov. 2008.
- [40] M. Wenk, *MIMO-OFDM-testbed: challenges, implementations, and measurement results*. ETH Zurich, 2010.
- [41] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Hardware impairments in large-scale MISO systems: Energy efficiency, estimation, and capacity limits," in *Proc. 8th Intern. Conf. Digital Signal Process.*, July 2013, pp. 1–6.
- [42] P. Zetterberg, "Experimental investigation of TDD reciprocity-based zero-forcing transmit precoding," *EURASIP J. Adv. Signal Process.*, vol. 2011, p. 5, Dec. 2011.
- [43] E. Björnson, P. Zetterberg, M. Bengtsson, and B. Ottersten, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, Jan. 2013.
- [44] K. Papadopoulos, *Channel Estimation for Large MIMO Systems Under Hardware Impairments*. McGill University Libraries, 2016.
- [45] A.-A. A. Boulogeorgos, N. D. Chatzidiamantis, and G. K. Karagiannidis, "Energy detection spectrum sensing under RF imperfections," *IEEE Trans. Commun.*, vol. 64, no. 7, pp. 2754–2766, July 2016.
- [46] E. Björnson, M. Matthaiou, and M. Debbah, "A new look at dual-hop relaying: Performance limits with hardware impairments," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4512–4525, Nov. 2013.
- [47] H. Suzuki, T. V. A. Tran, I. B. Collings, G. Daniels, and M. Hedley, "Transmitter noise effect on the performance of a MIMO-OFDM hardware implementation achieving improved coverage," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, Aug. 2008.
- [48] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *Int. ITG Workshop on Smart Antennas (WSA)*, IEEE, Feb. 2010, pp. 189–196.
- [49] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Sep. 2014.
- [50] M. Duarte, *Full-duplex wireless: Design, implementation and characterization*. Rice University, 2012.
- [51] M. Duarte and A. Sabharwal, "Full-duplex wireless communications using off-the-shelf radios: Feasibility and first results," in *Signals, Systems and Computers (ASILOMAR), 2010 Conference Record of the Forty Fourth Asilomar Conference on*. IEEE, 2010, pp. 1558–1562.

- [52] 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, Sept. 2014.
- [53] 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.
- [54] X. Li, C. Tepedelenlioğlu, and H. Şenol, "Channel estimation for residual self-interference in full-duplex amplify-and-forward two-way relays," *IEEE Trans. Wireless Commun.*, vol. 16, no. 8, pp. 4970–4983, Aug. 2017.
- [55] T. M. Kim and A. Paulraj, "Outage probability of amplify-and-forward cooperation with full duplex relay," in *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*. Paris, France: IEEE, Apr. 2012, pp. 75–79.
- [56] X. Li, C. Tepedelenlioğlu, and H. Şenol, "Optimal training for residual self-interference for full duplex one-way relays," *arXiv preprint arXiv:1709.06140*, 2017.
- [57] N. H. Tran, L. J. Rodríguez, and T. Le-Ngoc, "Optimal power control and error performance for full-duplex dual-hop AF relaying under residual self-interference," *IEEE Commun. Lett.*, vol. 19, no. 2, pp. 291–294, Feb. 2015.
- [58] G. Zheng, "Joint beamforming optimization and power control for full-duplex MIMO two-way relay channel," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 555–566, Feb. 2015.
- [59] 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.
- [60] T. Kwon, S. Lim, S. Choi, and D. Hong, "Optimal duplex mode for DF relay in terms of the outage probability," *IEEE Trans. Veh. Technol.*, vol. 59, no. 7, pp. 3628–3634, Sep. 2010.
- [61] S. Boyd, "Sequential convex programming," *Lecture Notes, Stanford University*, July 2008.
- [62] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.