Wideband Self-Interference Cancellation Beamforming with Digital Phased Arrays

Jordi Borras, *Member, IEEE*, and Roberto López-Valcarce, *Senior Member, IEEE*

***Abstract*—Simultaneous Transmit and Receive (STAR) systems face a significant implementation challenge in the form of self-interference (SI), which can saturate the Radio Frequency (RF) components of the co-located receive array and substantially degrade its performance. Digital transmit beamforming has been recently proposed as an effective means to deal with SI in narrowband settings. We explore digital precoding-based SI cancellation for wideband STAR based on Multiple-Input Multiple-Output Orthogonal Frequency-Division Multiplexing (MIMO-OFDM) under limited transmitter dynamic range. We propose two precoder designs respectively based on minimizing the worst-case SI or the total SI over the receive array, under a mutual information (MI) requirement. The non-convexity of both optimization problems is addressed through semidefinite relaxation, and in this way the optimal precoding matrices and number of data streams per subcarrier are obtained. For cases in which such number of streams cannot be supported, two suboptimal approaches based on power allocation and successive orthogonalization are proposed. Simulation results illustrate the SI reduction capability of the proposed precoder designs and highlight an MI-SI tradeoff.**

***Index Terms*—Simultaneous transmit and receive, digital transmit beamforming, self-interference cancellation, in-band full-duplex.**

I. Introduction

SIMULTANEOUS Transmit and Receive (STAR) has become a paradigm-shifting idea in wireless systems, with the potential not only to double the spectral efficiency through In-Band Full-Duplex (IBFD) communication [1] but also to enable several applications, *e.g.*, the integration of sensing and communications [2], covert communications [3], or joint opportunistic communication and spectrum sensing in interweave cognitive radio [4]. Despite its many advantages, dealing with the self-interference (SI) produced by direct leakage of the transmitted signal into the co-located receiver poses a significant challenge for practical deployments.

Recently, the pursuit of SI cancellation (SIC) schemes has led the research in the STAR landscape. In general terms, SIC can be performed by appropriately configuring the antenna interfaces [5–8], including the design of the optimal antenna separation, the radiation pattern of each element, or the polarization; implementing a dedicated analog SIC processing stage at the transceivers' radio-frequency (RF) front-end [9–12]; or incorporating the SIC capability into digital baseband processors [13–15]. Since the transmitter and the receiver are co-located, SI may completely overwhelm the latter, saturating its RF circuitry and/or analog-to-digital converters (ADC). RF SIC approaches are effective in this regard, as they address SI before saturation takes place, in contrast with digital SIC [16]. However, RF SIC schemes incur additional hardware and system complexity, degrade the receiver's noise figure, cannot avoid saturation of any preceding analog components, and scale poorly with the number of antennas, since each transmit-receive antenna pair requires a dedicated and programmable RF canceler. For these reasons, given the implementation advantages of digital processing, transmit beamforming has arisen as a promising SIC approach to mitigate SI in the wireless domain and provide enough transmitter-receiver isolation in multiantenna systems [17].

A. Related Works

Most existing works on Multiple- Input Multiple-Output Orthogonal Frequency-Division Multiplexing (MIMO-OFDM) STAR systems do not focus on minimizing SI under explicit Quality-of-Service (QoS) constraints; instead, they address the maximization of a communication figure of merit, such as the spectral efficiency (SE), while dealing with SI through additional mechanisms [18], *e.g.*, dedicated analog or digital SIC stages, or combining both [19]; see, for instance, [20–22].

When transmit beamforming is considered to mitigate SI without relying on these additional hardware-intensive stages, conventional approaches are typically based on spatial nulling [23–27], *i.e.,* projecting the transmitted signal onto the subspace spanned by the right singular vectors associated with the smallest (possibly zero) singular values of the SI channel matrix. This requires that the number of transmit (TX) antennas be larger than the number of co-located receive (RX) antennas. Some degradation in terms of beamforming gain is incurred, as a fraction of the available spatial degrees of freedom (DoF) is devoted to mitigating SI [27]. For these reasons, a number of techniques have been proposed which adjust the beamforming gain–SIC tradeoff. In [28–30] the TX beamformer is allowed to leak a certain amount of SI to be dealt with by means of dedicated mechanisms, such as digital or analog cancelers. But digital SIC cannot protect RF components from saturation, and the use of RF SIC brings back the scalability issue. A second alternative is to include the SI into the noise-plus-interference covariance matrix and maximizing the communication QoS [31–34]; however, this cannot generally avoid RF frontend

Work funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU under grants PID2022-136512OB-C21/C22, and by grant 2021 SGR 01033.

J. Borras was with the atlanTTic Research Center, Universidade de Vigo, 36310 Vigo, Spain. He is now with the Department of Signal Theory and Communications, Universitat Politècnica de Catalunya - BarcelonaTech (UPC), 08034 Barcelona, Spain. E-mail: jordi.borras@ieee.org

R. López-Valcarce is with the atlanTTic Research Center, Universidade de Vigo, 36310 Vigo, Spain. E-mail: valcarce@gts.uvigo.es.

saturation, since SI at the co-located RX array is not explicitly controlled. As a third alternative, communication QoS can be maximized under a constraint on the maximum admissible SI at the co-located RX array [35–38]. Although this approach allows to balance the beamforming gain–SIC tradeoff directly, setting the SI threshold to prevent saturation is hardware-dependent, making its generalization cumbersome.

Closer to our work, [39] focuses on minimizing SI subject to a QoS-type constraint in a narrowband scenario. Specifically, the TX beamformer is designed for worst-case SI minimization while guaranteeing a specific TX power in a known direction, implicitly assuming single-stream transmission. In wideband settings, [39] suggests to apply the proposed narrowband solution independently for each subcarrier.

Regardless of the efforts to mitigate SI at the TX side, hardware impairments due to the limited dynamic range of RF components (digital-to-analog converter (DAC), power amplifier, mixer) further aggravate the impact of SI in practice [40], [41]. This effect is generally overlooked in TX beamforming designs; or, if considered, SI-induced front-end saturation may not be explicitly addressed [42–45]. Whereas a few recent works [29], [32], [33], [35], [39] constitute the exception, they focus on narrowband settings. Focusing on [39], the impact of the limited dynamic range is not included in the adopted cost function, which is a significant limitation, since such impact typically depends on the per-antenna transmit power.

B. Contributions

We pursue the minimization of SI power at the receive antenna elements, accounting for the impact of limited transmitter dynamic range, under a QoS requirement in terms of Mutual Information (MI) with respect to the intended receiver. Our contributions can be summarized as follows:

- Two digital SIC precoder design problems are proposed in a MIMO-OFDM setting, respectively based on the minimization of worst-case SI and total SI at the input of the co-located receive array, with the purpose of avoiding front-end saturation. In contrast to previous works [39] that suggest addressing the wideband case in a *per-subcarrier* (PS) fashion, the precoding matrices for each subcarrier are jointly designed in this work, resulting in significantly improved performance.
- Both optimization problems are particular instances of the non-convex Generalized Quadratic Matrix Program (GQPM) [46]. Resorting to the Semidefinite Relaxation (SDR) framework [47], we study the joint design of the precoding matrices and the number of data streams per subcarrier. The SDR approach is shown to be tight for the problems at hand. Cases in which the resulting number of streams cannot be supported by the transmitter are handled through suboptimal schemes based on power allocation and successive orthogonalization.
- We take into account the effect of limited dynamic range of the transmit array RF circuitry in the cost function. Since the power of the transmit noise induced by hardware impairments is proportional to the per-antenna transmit power [40], this general approach incurs an additional per-antenna power penalty.

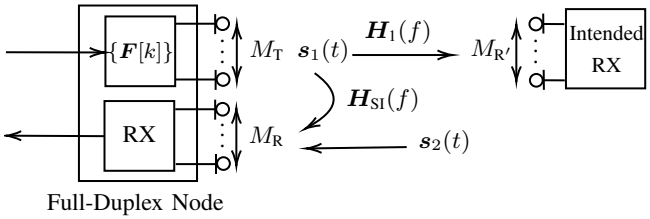


Fig. 1: MIMO-OFDM scenario with a Full-Duplex (FD) node simultaneously transmitting to an intended RX and receiving a signal from the environment.

- Numerical results illustrate the MI-SI tradeoff and the incurred QoS degradation when SI is maximally canceled through digital beamforming. Moreover, the advantages with respect to the PS approach are also highlighted.

C. Organization and Notation

In Sec. II the system model is presented. The general design of SIC precoding matrices is addressed in Sec. III. Sec. IV describes the design procedure under a fixed number of data streams. Numerical results are reported in Sec. V, and conclusions are drawn in Sec. VI.

Notation: We use boldface lowercase (resp., uppercase) symbols to denote vectors (resp., matrices). $\boldsymbol{A}^T$, $\boldsymbol{A}^H$, $\boldsymbol{A}^\sharp$, $\mathrm{tr}(\boldsymbol{A})$, $\|\cdot\|_{\mathrm{F}}$, and $|\boldsymbol{A}|$ denote the transpose, transpose conjugate, pseudoinverse, trace, Frobenius norm, and determinant of matrix $\boldsymbol{A}$, respectively. $\boldsymbol{I}_K$ is the $K \times K$ identity matrix, and its i-th column is denoted as $\boldsymbol{e}_i$. $\boldsymbol{A} \succeq \boldsymbol{0}$ indicates that matrix $\boldsymbol{A}$ is positive semidefinite, and $\mathrm{diag}(\boldsymbol{A})$ returns a diagonal matrix with the main diagonal of the square matrix $\boldsymbol{A}$. $\mathbb{H}^m$ and $\mathbb{D}^m$ stand for the sets of $m \times m$ Hermitian matrices and $m \times m$ diagonal matrices, respectively. $\delta(\cdot)$ is the Dirac delta function, $(x)^+ = \max(0, x)$, and $\mathbb{1}_{(\cdot)}$ is the indicator function, which equals 1 if the condition $(\cdot)$ holds and 0 otherwise. The circular complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ is denoted as $\mathcal{N}_{\mathbb{C}}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. $\mathbb{E}\{\cdot\}$ stands for statistical expectation.

II. System Model

Consider the MIMO-OFDM setting of Fig. 1, in which an FD node with separate transmit-receive arrays simultaneously transmits a signal $\boldsymbol{s}_1(t)$ to a multi-antenna node and receives a signal of interest $\boldsymbol{s}_2(t)$. The nature of $\boldsymbol{s}_2(t)$ is not relevant to our study, and thus is left unspecified. The FD node is equipped with M_{T} TX and M_{R} RX antennas, whereas the intended RX has $M_{\mathrm{R}'}$ antennas. The frequency response of the channel from the FD node to the intended RX is denoted as $\boldsymbol{H}_1(f) \in \mathbb{C}^{M_{\mathrm{R}'} \times M_{\mathrm{T}}}$. Analogously, the frequency response of the SI channel between the TX and RX arrays of the FD node is denoted as $\boldsymbol{H}_{\mathrm{SI}}(f) \in \mathbb{C}^{M_{\mathrm{R}} \times M_{\mathrm{T}}}$. As in, *e.g.*, [28], [48–50], both $\boldsymbol{H}_1(f)$ and $\boldsymbol{H}_{\mathrm{SI}}(f)$ are assumed known, in order to explore the theoretical limits of SIC in MIMO-OFDM settings. In practice, the channel responses can be estimated in an initial training stage. In wideband FD settings, several strategies have been proposed in the literature. For instance, [51] exploits the fact that the transmitted symbols are known at the co-located FD receive array and develops a maximum likelihood SI channel estimator. A computationally simpler approach is to exploit the sparsity of the SI channel, leveraging compressed sensing

techniques [52]. From a different perspective, [53] proposes the use of orthogonal pilot sequences to jointly estimate SI and communication (intended) channels. Our results provide upper limits to the performance of practical systems which invariably will incorporate channel estimation errors.

In the next subsections we provide a rigorous characterization of the statistics of the transmit signal $\boldsymbol{s}_1(t)$, of the received signal at the intended receiver, and of the self-interference generated at the receive array of the FD node, respectively.

A. Transmit Signal

The FD node uses K subcarriers, with indices comprised in $\mathcal{K} = \{k_1, k_2, \ldots, k_K\}$, to send M OFDM symbols denoted by $\boldsymbol{z}_m[k] \in \mathbb{C}^{M_{\mathrm{S}}[k]}$, where $k \in \mathcal{K}$ and $m \in \{1, \ldots, M\}$ are the subcarrier and symbol indices, respectively, and

$$M_{\mathrm{S}}[k] \leq d \triangleq \min(M_{\mathrm{T}}, M_{\mathrm{R}'}) \tag{1}$$

is the number of data streams per subcarrier.

The data are assumed zero-mean identically distributed and uncorrelated across OFDM symbols and subcarriers, so that $\mathbb{E}\{\boldsymbol{z}_m[k]\boldsymbol{z}_{m'}^H[k']\} = \frac{P_{\mathrm{T}}}{KM_{\mathrm{S}}[k]}\boldsymbol{I}_{M_{\mathrm{S}}[k]}$ if $m = m'$ and $k = k'$, where P_{T} is the total transmit power; otherwise, $\mathbb{E}\{\boldsymbol{z}_m[k]\boldsymbol{z}_{m'}^H[k']\} = \boldsymbol{0}_{M_{\mathrm{S}}[k]\times M_{\mathrm{S}}[k']}$.

Let $\boldsymbol{F}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{S}}[k]}$ be the precoding matrix at the k-th subcarrier, so that $\boldsymbol{x}_m[k] = \boldsymbol{F}[k]\boldsymbol{z}_m[k] \in \mathbb{C}^{M_{\mathrm{T}}}$ is the vector modulated onto the k-th subcarrier in the m-th OFDM symbol. The FD node employs M_{T}-parallel N-point IDFTs such that the discrete-time baseband transmitted signal reads as

$$\widetilde{\boldsymbol{s}}_1[n] = \sum_{m=1}^{M}\sum_{k\in\mathcal{K}} \boldsymbol{x}_m[k]p[n-mL]e^{\imath\frac{2\pi k(n-mL)}{N}}, \tag{2}$$

where $L = N + N_{\mathrm{CP}}$, with N_{CP} the Cyclic Prefix (CP) length in samples, and $p[n]$ is the shaping pulse with Fourier transform $P(e^{\imath\omega}) = \sum_n p[n]e^{-\imath\omega n}$. Using a DAC interpolation filter $h_{\mathrm{I}}(t)$, with frequency response $H_{\mathrm{I}}(f)$, the baseband continuous-time transmitted signal is given by

$$\boldsymbol{s}_1(t) = \widetilde{\boldsymbol{s}}_1(t) + \boldsymbol{n}_{\mathrm{T}}(t) = \sum_{n=-\infty}^{\infty} \widetilde{\boldsymbol{s}}_1[n]h_{\mathrm{I}}(t-nT_{\mathrm{s}}) + \boldsymbol{n}_{\mathrm{T}}(t), \tag{3}$$

with T_{s} the sampling period (thus, $\Delta_f = \frac{1}{NT_{\mathrm{s}}}$ is the subcarrier spacing). The TX noise vector $\boldsymbol{n}_{\mathrm{T}}(t) \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\mathrm{T}})$ models the transmitter's limited dynamic range [40]. As reported in [40], [43–45], [53], experimental studies corroborate that the independent Gaussian noise model for the transmit noise $\boldsymbol{n}_{\mathrm{T}}(t)$ approximates well the combination of several hardware impairments[1], including the effects of additive power amplifier noise, nonlinearities in the DAC and power amplifier, and mixer phase noise [54–57].

Let $\phi_0(f) = P^*(e^{\imath 2\pi f T_{\mathrm{s}}})$, and $\phi_k(f) = \phi_0(f - k\Delta_f)$. For instance, with a rectangular pulse $p[n] = 1$ for $0 \leq n \leq L-1$ and zero otherwise, $\phi_0(f) = \frac{\mathrm{sinc}(LT_{\mathrm{s}}f)}{\mathrm{sinc}(T_{\mathrm{s}}f)}Le^{\imath\pi(L-1)T_{\mathrm{s}}f}$. Following [58], the power spectral density (PSD) matrix of the signal component $\widetilde{\boldsymbol{s}}_1(t)$ in (3) can be shown to be

$$\boldsymbol{S}_{\widetilde{\boldsymbol{s}}_1}(f) = \frac{P_{\mathrm{T}}}{K}\frac{|H_{\mathrm{I}}(f)|^2}{LT_{\mathrm{s}}}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}|\phi_k(f)|^2\boldsymbol{F}[k]\boldsymbol{F}^H[k]. \tag{4}$$

Let us denote the integrated response of the k-th subcarrier as

$$\alpha_k^2 \triangleq \frac{1}{LT_{\mathrm{s}}}\int_{-\infty}^{\infty}|H_{\mathrm{I}}(f)|^2|\phi_k(f)|^2\mathrm{d}f. \tag{5}$$

Note, $H_{\mathrm{I}}(f)$ is lowpass with cutoff frequency $\frac{1}{2T_{\mathrm{s}}}$. If the number of subcarriers K is large, then the response of each subcarrier, $H_{\mathrm{I}}(f)\phi_k(f)$, will tend to a narrow spike centered at $f = k\Delta_f$, and it becomes reasonable to approximate

$$\varphi_k(f) \triangleq \frac{1}{LT_{\mathrm{s}}}|H_{\mathrm{I}}(f)|^2|\phi_k(f)|^2 \approx \alpha_k^2\delta(f - k\Delta f). \tag{6}$$

In addition, if the DAC filter response $H_{\mathrm{I}}(f)$ is sufficiently flat within its passband, it is reasonable to assume that α_k^2 is in fact independent of k; thus, without loss of generality, we will take $\alpha_k^2 = 1$ for all $k \in \mathcal{K}$. With this, the power delivered to the ℓ-th transmit antenna, $\ell = 1, \ldots, M_{\mathrm{T}}$, becomes

$$\begin{aligned} P_{\mathrm{T},\ell} &\triangleq \int_{-\infty}^{\infty}\boldsymbol{e}_\ell^H\boldsymbol{S}_{\widetilde{\boldsymbol{s}}_1}(f)\boldsymbol{e}_\ell\mathrm{d}f \\ &= \frac{P_{\mathrm{T}}}{K}\boldsymbol{e}_\ell^H\left(\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\boldsymbol{F}[k]\boldsymbol{F}^H[k]\right)\boldsymbol{e}_\ell. \end{aligned} \tag{7}$$

As per [40, Sec. II-C], the TX noise $\boldsymbol{n}_{\mathrm{T}}(t)$ is approximately temporally white, independent of the signal $\widetilde{\boldsymbol{s}}_1(t)$; its covariance matrix $\boldsymbol{\Sigma}_{\mathrm{T}}$ is diagonal, and the variance of its ℓ-th component is proportional to the power of the useful signal sent through the corresponding antenna, *i.e.,* to $P_{\mathrm{T},\ell}$. Thus,

$$\boldsymbol{\Sigma}_{\mathrm{T}} = \frac{1}{\eta_{\mathrm{T}}}\frac{P_{\mathrm{T}}}{K}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\mathrm{diag}\left(\boldsymbol{F}[k]\boldsymbol{F}^H[k]\right), \tag{8}$$

with η_{T} the Signal-to-Noise Ratio (SNR) at each TX antenna.

B. Received Signal

In the absence of synchronization errors, the signal reaching the intended RX at the k-th subcarrier in the m-th OFDM symbol, after CP removal and DFT processing, is given by

$$\boldsymbol{y}_m[k] = \boldsymbol{H}_1[k]\left(\boldsymbol{F}[k]\boldsymbol{z}_m[k] + \boldsymbol{n}_{\mathrm{T}}[k]\right) + \boldsymbol{n}[k], \tag{9}$$

with $\boldsymbol{H}_1[k] = \boldsymbol{H}_1(f)|_{f=k\Delta_f} \in \mathbb{C}^{M_{\mathrm{R}'}\times M_{\mathrm{T}}}$ the intended channel matrix at the k-th subcarrier, $\boldsymbol{n}[k] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \sigma^2\boldsymbol{I}_{M_{\mathrm{R}'}})$ the additive Gaussian thermal noise, and $\boldsymbol{n}_{\mathrm{T}}[k]$ the frequency-domain TX distortion at the k-th subcarrier, whose covariance matrix is still given by (8) [43, Lemma II.1]. Hence, the global noise term at the intended RX, $\boldsymbol{\nu}[k] = \boldsymbol{H}_1[k]\boldsymbol{n}_{\mathrm{T}}[k] + \boldsymbol{n}[k]$, can be statistically modeled as $\boldsymbol{\nu}[k] \sim \mathcal{N}_{\mathbb{C}}(\boldsymbol{0}, \boldsymbol{\Sigma}_{\boldsymbol{\nu}}[k])$, with

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\nu}}[k] &= \boldsymbol{H}_1[k]\boldsymbol{\Sigma}_{\mathrm{T}}\boldsymbol{H}_1^H[k] + \sigma^2\boldsymbol{I}_{M_{\mathrm{R}'}} \\ &\approx \sigma^2\boldsymbol{I}_{M_{\mathrm{R}'}}, \end{aligned} \tag{10}$$

where the approximation is due to the fact that, in practice, the effect of thermal noise generated locally at the intended receiver is much larger than that of the term due to transmit noise, since the SNR per transmit antenna η_{T} is usually large, and this transmit noise term is affected by channel attenuation.

[1]Hardware impairments introduce complex phenomena such as intercarrier leakage and spectral regrowth. In-band intercarrier leakage is the frequency-domain manifestation of time-domain distortion; since the receiver's DFT computes each subcarrier as the weighted summation of time-domain samples, the Central Limit Theorem ensures this aggregate leakage converges to a circularly symmetric complex Gaussian noise. Meanwhile, spectral regrowth produces out-of-band emissions that are mitigated by RF filtering and digital pre-distortion, falling outside the scope of this work.

C. Self-Interference

Transmission of $\boldsymbol{s}_1(t)$ causes SI in the FD receive array; if unchecked, this SI may saturate the receive array RF chains. Let $\boldsymbol{S}_{\boldsymbol{s}_1}(f) = \boldsymbol{S}_{\widetilde{\boldsymbol{s}}_1}(f) + \boldsymbol{\Sigma}_{\mathrm{T}} \in \mathbb{C}^{M_{\mathrm{T}} \times M_{\mathrm{T}}}$ be the PSD matrix of the TX signal $\boldsymbol{s}_1(t)$. Then, the SI power induced on the i-th co-located receive antenna is given by

$$P_{\mathrm{SI},i} \triangleq \int_{-\infty}^{\infty} \boldsymbol{e}_i^H \boldsymbol{H}_{\mathrm{SI}}(f) \boldsymbol{S}_{\boldsymbol{s}_1}(f) \boldsymbol{H}_{\mathrm{SI}}^H(f) \boldsymbol{e}_i \mathrm{d}f, \tag{11}$$

whose explicit form in a MIMO-OFDM STAR system is characterized in the following lemma. First, let us introduce the matrices $\boldsymbol{C}_i[k] \in \mathbb{C}^{M_{\mathrm{T}} \times M_{\mathrm{T}}}$ as

$$\begin{aligned} \boldsymbol{C}_i[k] &\triangleq \int_{-\infty}^{\infty} \boldsymbol{H}_{\mathrm{SI}}^H(f) \boldsymbol{e}_i \boldsymbol{e}_i^H \boldsymbol{H}_{\mathrm{SI}}(f) \varphi_k(f) \mathrm{d}f & (12) \\ &\approx \boldsymbol{H}_{\mathrm{SI}}^H[k] \boldsymbol{e}_i \boldsymbol{e}_i^H \boldsymbol{H}_{\mathrm{SI}}[k], & (13) \end{aligned}$$

where $\boldsymbol{H}_{\mathrm{SI}}[k] \triangleq \boldsymbol{H}_{\mathrm{SI}}(f)|_{f=k\Delta_f}$, and we have used (6) with $\alpha_k^2 = 1$.

Lemma 1. *Let $g_{ij} \triangleq \int_{-\infty}^{\infty} |\boldsymbol{e}_i^H \boldsymbol{H}_{\mathrm{SI}}(f) \boldsymbol{e}_j|^2 \mathrm{d}f$ be the power gain of the SI channel from the j-th TX to i-th RX antennas of the FD node. The SI power at the i-th receive antenna given by* (11) *amounts to*

$$P_{\mathrm{SI,i}} = \frac{P_{\mathrm{T}}}{K} \mathrm{tr} \sum_{k \in \mathcal{K}} \frac{1}{M_S[k]} \boldsymbol{F}^H[k] \underbrace{\left(\boldsymbol{C}_i[k] + \sum_{j=1}^{M_T} \frac{g_{ij}}{\eta_T} \boldsymbol{e}_j \boldsymbol{e}_j^H \right)}_{\triangleq \widetilde{\boldsymbol{C}}_i[k]} \boldsymbol{F}[k]. \tag{14}$$

Proof. See Appendix A. □

Note that the SI component due to TX noise is not neglected, since in general the attenuation of the SI channel is not sufficient to bring such component down to the level of the FD receiver's thermal noise.

In the framework of the described model, we address the design of the precoding matrices $\{\boldsymbol{F}[k]\}_{k \in \mathcal{K}}$ to minimize the SI induced on the co-located RX array while meeting a QoS constraint on the intended communication link. The design of combining matrices at both intended and co-located receivers is out of the scope of this work. Once $\{\boldsymbol{F}[k]\}_{k \in \mathcal{K}}$ are found, the intended receiver may design its combiner following any standard approach, *e.g.,* to maximize SE. On the other hand, combiner design at the co-located receiver is application-dependent, *e.g.*, maximize SE in IBFD or maximize the radar Signal-to-Interference-plus-Noise Ratio (SINR) in Integrated Sensing and Communications (ISAC).

III. General Approach: Variable Number of Data Streams

Having established the system model and characterized the statistics of the signals involved, in this section we present our precoder designs in their more general form, in which the number of data streams may differ across subcarriers. We state the mathematical problems with their corresponding QoS constraints, present a suitable semidefinite relaxation (SDR) approach, and discuss its computational complexity.

A. Problem Statement

We first approach the design of the precoders from a broad perspective, assuming that the number of data streams per subcarrier, $\{M_{\mathrm{S}}[k]\}_{k \in \mathcal{K}}$, are not fixed beforehand; rather, they will be optimized to maximally suppress SI while meeting a QoS constraint on the intended link. The adopted QoS metric is the MI, which under Gaussian signaling reads as

$$\begin{aligned} \mathcal{I} &= \frac{1}{K} \sum_{k \in \mathcal{K}} \log \Big| \boldsymbol{I}_{M_{\mathrm{R}'}} \\ &\quad + \frac{P_{\mathrm{T}}}{K M_{\mathrm{S}}[k]} \boldsymbol{\Sigma}_{\boldsymbol{\nu}}^{-1}[k] \boldsymbol{H}_1[k] \boldsymbol{F}[k] \boldsymbol{F}^H[k] \boldsymbol{H}_1^H[k] \Big| \\ &\approx \frac{1}{K} \sum_{k \in \mathcal{K}} \log \left| \boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma_k \boldsymbol{H}_1[k] \boldsymbol{F}[k] \boldsymbol{F}^H[k] \boldsymbol{H}_1^H[k] \right|, \end{aligned} \tag{15}$$

where $\gamma_k \triangleq \frac{P_{\mathrm{T}}}{K M_{\mathrm{S}}[k] \sigma^2}$ is the (per-stream) channel SNR [40] at subcarrier k, and the approximation in (10) has been used. The dependence of $\mathcal{I}$ on the design parameters $M_{\mathrm{S}}[k]$ (via γ_k) makes (15) difficult to handle. To sidestep this issue, let $\gamma \triangleq \frac{P_{\mathrm{T}}}{K d \sigma^2}$ and note from (1) that $\gamma \leq \gamma_k$, $\forall k$. Therefore, the alternative QoS metric

$$\widetilde{\mathcal{I}} \triangleq \frac{1}{K} \sum_{k \in \mathcal{K}} \log \left| \boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma \boldsymbol{H}_1[k] \boldsymbol{F}[k] \boldsymbol{F}^H[k] \boldsymbol{H}_1^H[k] \right|, \tag{16}$$

provides a lower bound on $\mathcal{I}$, and will be used in the sequel to set the QoS constraint.

Let $q(\boldsymbol{p}_{\mathrm{SI}})$, with $\boldsymbol{p}_{\mathrm{SI}} \triangleq [\, P_{\mathrm{SI},1} \cdots P_{\mathrm{SI},M_{\mathrm{R}}} \,]^T$, be some function of the SI power levels (14) at all co-located RX antennas, penalizing large SI values. The proposed design of the precoders $\{\boldsymbol{F}[k]\}_{k \in \mathcal{K}}$ is cast as

$$\begin{aligned} \min_{\{\boldsymbol{F}[k]\}_{k \in \mathcal{K}}} \quad & q(\boldsymbol{p}_{\mathrm{SI}}) & (17a) \\ \text{s.t.} \quad & S_{\mathrm{T}} \triangleq \mathrm{tr} \int_{-\infty}^{\infty} \boldsymbol{S}_{\widetilde{\boldsymbol{s}}_1}(f) \mathrm{d}f \leq P_{\mathrm{T}}, & (17b) \\ & \widetilde{\mathcal{I}} \geq t, & (17c) \end{aligned}$$

where (17b) limits the total TX power and (17c) is the MI constraint. We shall consider two choices for $q(\boldsymbol{p}_{\mathrm{SI}})$:

$$\begin{aligned} q_{\mathrm{sum}} &= \mathrm{tr} \sum_{i=1}^{M_{\mathrm{R}}} \sum_{k \in \mathcal{K}} \boldsymbol{F}^H[k] \widetilde{\boldsymbol{C}}_i[k] \boldsymbol{F}[k], & (18) \\ q_{\max} &= \max_{i \in \{1, \ldots, M_{\mathrm{R}}\}} \mathrm{tr} \sum_{k \in \mathcal{K}} \boldsymbol{F}^H[k] \widetilde{\boldsymbol{C}}_i[k] \boldsymbol{F}[k]. & (19) \end{aligned}$$

By setting $q(\boldsymbol{p}_{\mathrm{SI}}) = q_{\mathrm{sum}}$ one aims at minimizing the total SI induced on the FD receiver, whereas with $q(\boldsymbol{p}_{\mathrm{SI}}) = q_{\max}$ the goal is to minimize the worst-case SI over all antennas.

Substituting (4) in the expression of S_{T} in (17b), and using approximation (6) with $\alpha_k^2 = 1$ $\forall k$, it is readily found that

$$S_{\mathrm{T}} \approx \frac{P_{\mathrm{T}}}{K} \sum_{k \in \mathcal{K}} \frac{1}{M_{\mathrm{S}}[k]} \|\boldsymbol{F}[k]\|_{\mathrm{F}}^2. \tag{20}$$

Hence, the power constraint $S_{\mathrm{T}} \leq P_{\mathrm{T}}$ in (17b) amounts to

$$\sum_{k \in \mathcal{K}} \frac{1}{M_{\mathrm{S}}[k]} \|\boldsymbol{F}[k]\|_F^2 \leq K. \tag{21}$$

Since $M_{\mathrm{S}}[k] \geq 1$, $\forall k \in \mathcal{K}$, note that $\sum_{k\in\mathcal{K}} \frac{1}{M_{\mathrm{S}}[k]} \|\boldsymbol{F}[k]\|_F^2 \leq \sum_{k\in\mathcal{K}} \|\boldsymbol{F}[k]\|_F^2$. To make the ensuing problems tractable, we replace (21) with the tighter constraint

$$\sum_{k\in\mathcal{K}} \|\boldsymbol{F}[k]\|_F^2 \leq K. \tag{22}$$

Irrespective of the chosen criterion (18) or (19), design (17) is a particular instance of the so-called Generalized Quadratic Matrix Programming (GQMP) problem [46], which is non-convex. By exploiting the cyclic property of the trace, the objective functions and constraints can be written in terms of the positive semidefinite matrices $\boldsymbol{X}[k] = \boldsymbol{F}[k]\boldsymbol{F}^H[k] \in \mathbb{H}^{M_{\mathrm{T}}}$, $\forall k \in \mathcal{K}$. Note that $\mathrm{rank}(\boldsymbol{X}[k]) \leq M_{\mathrm{S}}[k] \leq d$, where d is the number of available spatial DoF, see (1); and that the MI (16) can be written as

$$\widetilde{\mathcal{I}} = \frac{1}{K} \sum_{k\in\mathcal{K}} \log \left| \boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma \boldsymbol{H}_1[k] \boldsymbol{X}[k] \boldsymbol{H}_1^H[k] \right|, \tag{23}$$

which is concave in $\boldsymbol{X}[k]$, $\forall k \in \mathcal{K}$. Then, on the one hand, when $q(\boldsymbol{p}_{\mathrm{SI}}) = q_{\mathrm{sum}}$, problem (17) becomes

$$\begin{aligned}
\mathbb{P}_1: \quad \min_{\{\boldsymbol{X}[k]\in\mathbb{H}^{M_{\mathrm{T}}}\}} \quad & q_{\mathrm{sum}} && \text{(24a)}\\
\text{s.t.} \quad & \mathrm{tr} \sum_{k\in\mathcal{K}} \boldsymbol{X}[k] \leq K && \text{(24b)}\\
& \widetilde{\mathcal{I}} \geq t && \text{(24c)}\\
& \boldsymbol{X}[k] \succeq \boldsymbol{0},\ \forall k \in \mathcal{K} && \text{(24d)}\\
& \mathrm{rank}(\boldsymbol{X}[k]) \leq d,\ \forall k \in \mathcal{K}. && \text{(24e)}
\end{aligned}$$

On the other hand, with $q(\boldsymbol{p}_{\mathrm{SI}}) = q_{\max}$, (17) reads as

$$\begin{aligned}
\mathbb{P}_2: \quad \min_{\{\boldsymbol{X}[k]\in\mathbb{H}^{M_{\mathrm{T}}}\}} \quad & \beta && \text{(25a)}\\
\text{s.t.} \quad & \mathrm{tr} \sum_{k\in\mathcal{K}} \widetilde{\boldsymbol{C}}_i[k] \boldsymbol{X}[k] \leq \beta \ \ \forall i && \text{(25b)}\\
& \text{(24b), (24c), (24d), (24e).} && \text{(25c)}
\end{aligned}$$

At the solution of either $\mathbb{P}_1$ or $\mathbb{P}_2$, the MI constraint must be met with equality: $\widetilde{\mathcal{I}} = t$. Otherwise, it would be possible to scale the corresponding matrices $\{\boldsymbol{X}[k]\}$ by $1-\epsilon$ for some small $\epsilon > 0$, in such a way that all the constraints are satisfied while the objective is reduced. On the other hand, the transmit power constraint may or may not be met with equality, depending on how stringent the MI constraint is.

B. A Remark on the QoS constraint

The following consideration can be made about the MI requirement (17c). For $s \in \{1, \ldots, d\}$, let us define

$$R(s) \triangleq \max_{\{\boldsymbol{F}[k]\}} \tilde{\mathcal{I}} \quad \text{s.t.} \begin{cases} \sum_{k\in\mathcal{K}} \|\boldsymbol{F}[k]\|_F^2 & \leq & K, \\ \mathrm{rank}(\boldsymbol{F}[k]) & = & s,\ \forall k \in \mathcal{K}. \end{cases} \tag{26}$$

Thus, $R(s)$ corresponds to the maximum MI attainable in the intended channel under the power constraint $S_{\mathrm{T}} \leq P_{\mathrm{T}}$ when s data streams are transmitted on each subcarrier, and is obtained as $R(s) = \frac{1}{K} \sum_{k\in\mathcal{K}} \sum_{i=1}^{s} \log\left(1 + \sigma_i^2[k]\lambda_i[k]\right)$, with $\sigma_i[k]$ the i-th largest singular value of $\boldsymbol{H}_1[k]$, and $\lambda_i[k]$ the associated water-filling coefficient. Clearly, the QoS requirement in (17c) should not exceed the maximum attainable MI, so we set $t = R(d)/\rho$, with $\rho \geq 1$ establishing the MI-SI tradeoff.

C. A Semidefinite Relaxation Approach

Neither $\mathbb{P}_1$ nor $\mathbb{P}_2$ are convex, due to the rank constraint on $\boldsymbol{X}[k]$. By dropping this constraint, $\mathbb{P}_1$ and $\mathbb{P}_2$ are relaxed to semidefinite convex problems [47], [59], [60], which can be solved numerically via standard interior-point method (IPM) schemes. An important issue regarding this SDR approach is whether the corresponding solutions are close to those of the original non-convex problems. Following [59], [60], we say that the SDR approach is *tight* for $\mathbb{P}_\ell$ whenever the corresponding SDR solution, say $\{\boldsymbol{X}_\ell^\star[k]\}_{k\in\mathcal{K}}$, satisfies $\mathrm{rank}(\boldsymbol{X}_\ell^\star[k]) \leq d$, $\forall k \in \mathcal{K}$. In this regard, the following holds.

Theorem 1. *The SDR approach is tight for $\mathbb{P}_\ell$, $\ell \in \{1, 2\}$. Therefore, it provides the true solution to problem $\mathbb{P}_\ell$.*

Proof. See Appendix B. □

Thus, the number of streams per subcarrier and the precoders can be found from the SDR solution $\{\boldsymbol{X}_\ell^\star[k]\}_{k\in\mathcal{K}}$ as

$$M_{\mathrm{S},\ell}^\star[k] = \mathrm{rank}(\boldsymbol{X}_\ell^\star[k]), \qquad \boldsymbol{F}_\ell^\star[k] = \widetilde{\boldsymbol{U}}_\ell[k] \widetilde{\boldsymbol{D}}_\ell^{\frac{1}{2}}[k], \tag{27}$$

where $\widetilde{\boldsymbol{U}}_\ell[k] \in \mathbb{C}^{M_{\mathrm{T}} \times M_{\mathrm{S},\ell}^\star[k]}$ contains the unit-norm eigenvectors of $\boldsymbol{X}_\ell^\star[k]$ associated with the $M_{\mathrm{S},\ell}^\star[k]$ largest eigenvalues, and $\widetilde{\boldsymbol{D}}_\ell[k]$ is diagonal with the corresponding eigenvalues.

D. On the Computational Complexity

The rank-relaxed versions of $\mathbb{P}_1$ and $\mathbb{P}_2$ are multi-block Semidefinite Programs (SDPs) with K Hermitian positive semidefinite variables $\{\boldsymbol{X}_k\}_{k\in\mathcal{K}}$ of size $M_{\mathrm{T}} \times M_{\mathrm{T}}$, which are coupled through the the convex, non-affine constraint (24c). Solving a general multi-block SDP with n Hermitian positive semidefinite blocks of size $m \times m$ via primal-dual IPMs with fixed accuracy ϵ requires at most $\mathcal{O}\left(\sqrt{mn}\log(\epsilon^{-1})\right)$ iterations [61], with a cost per iteration of $\mathcal{O}\left(nm^3\right) + \mathcal{O}\left(qnm^2\right) + \mathcal{O}\left(q^3\right)$, where q is the total number of affine constraints [62]. Since $\mathbb{P}_1$ and $\mathbb{P}_2$ respectively have one and $(M_{\mathrm{R}}+1)$ affine constraints, their worst-case computational cost is respectively $\mathcal{O}\left(KM_{\mathrm{T}}^3\right) + \mathcal{O}\left(KM_{\mathrm{T}}^2\right) + \mathcal{O}\left(1\right)$ and $\mathcal{O}\left(KM_{\mathrm{T}}^3\right) + \mathcal{O}\left((M_{\mathrm{R}}+1)KM_{\mathrm{T}}^2\right) + \mathcal{O}\left((M_{\mathrm{R}}+1)^3\right)$, each requiring at most $\mathcal{O}(\sqrt{KM_{\mathrm{T}}}\log(\epsilon^{-1}))$ IPM iterations. For comparison, spatial nulling techniques (see, *e.g.*, [23]), involve the computation of two singular value decompositions (SVDs) per subcarrier implying a complexity of $\mathcal{O}\left(K\max(M_{\mathrm{T}}, M_{\mathrm{R}})(\min(M_{\mathrm{T}}, M_{\mathrm{R}}))^2\right) + \mathcal{O}\left(K\max(M_{\mathrm{T}}, M_{\mathrm{R}'})(\min(M_{\mathrm{T}}, M_{\mathrm{R}'}))^2\right)$.

Solving the rank-relaxed versions of $\mathbb{P}_1$ and $\mathbb{P}_2$ through interior point methods can be costly for large M_{T} (the dimension of the matrices $\boldsymbol{X}[k]$). For $\mathbb{P}_1$, the problem admits an equivalent low-dimensional convex formulation:

Corollary 1. *For the case of $\mathbb{P}_1$, there exists an equivalent reduced form of the convex problem in which the optimization variables $\{\boldsymbol{Y}_{11}[k],\ k \in \mathcal{K}\}$ are of dimension $M_{R'} \times M_{R'}$.*

Proof. See Appendix B. □

This will result in significant computational savings whenever $M_{\mathrm{T}} \gg M_{\mathrm{R}'}$, *i.e.,* when the TX array of the full-duplex node (*e.g.,* a base station, BS) is much larger than the RX array of the intended receiver (*e.g.,* a mobile device).

IV. Fixing the Number of Data Streams

The SDR-based precoder design from Sec. III jointly optimizes the precoding matrices and the number of data streams per subcarrier. However, there may be settings in which the number of data streams is fixed beforehand. For instance, according to 3GPP 5G NR Release 18 [63], the number of streams sent per user-equipment is limited to 8. Thus, the SDR solution may not be admissible if $M^{\star}_{\mathrm{S},\ell}[k]$ in (27) turns out to be larger than some prescribed number. For this reason, in this section we present two alternative, suboptimal approaches suitable for scenarios with a fixed number of data streams per subcarrier, say $M_{\mathrm{S}}[k] = M_{\mathrm{S}}$, $\forall k \in \mathcal{K}$. We write the precoding matrices as $\boldsymbol{F}[k] = \boldsymbol{V}[k]\boldsymbol{\Lambda}^{\frac{1}{2}}[k]$, with $\boldsymbol{V}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{S}}}$ a semi-unitary matrix and $\boldsymbol{\Lambda}[k] \in \mathbb{R}^{M_{\mathrm{S}}\times M_{\mathrm{S}}}$ a positive semidefinite diagonal matrix; this is without loss of generality, since the problems depend on $\boldsymbol{F}[k]$ only through the product $\boldsymbol{F}[k]\boldsymbol{F}^H[k] = \boldsymbol{V}[k]\boldsymbol{\Lambda}[k]\boldsymbol{V}^H[k]$. We refer to $\{\boldsymbol{\Lambda}[k]\}$ as the *power allocation matrices.*

A. First Approach: Per-Subcarrier Power Allocation

Our first approach is based on truncating the rank of SDR solutions, retaining the corresponding singular vectors, and then optimizing the power allocation matrices. Specifically, we first solve the SDR version of $\mathbb{P}_\ell$, setting the MI threshold to $t = R(M_{\mathrm{S}})/\rho$, to account for the fixed number of data streams M_{S}. Letting $\{\boldsymbol{X}^{\star}_{\ell}[k]\}_{k\in\mathcal{K}}$ be the obtained SDR solutions, we then take $\boldsymbol{V}_\ell[k]$ as the M_{S} dominant eigenvectors of $\boldsymbol{X}^{\star}_{\ell}[k]$. With these, and to ease the notation, let us define at this point the following $M_{\mathrm{S}} \times M_{\mathrm{S}}$ matrices, respectively related to the effective intended and SI channels:

$$\boldsymbol{\Omega}_\ell[k] \triangleq \boldsymbol{V}_\ell^H[k]\boldsymbol{H}_1^H[k]\boldsymbol{H}_1[k]\boldsymbol{V}_\ell[k], \tag{28}$$

$$\boldsymbol{\Xi}_{i,\ell}[k] \triangleq \boldsymbol{V}_\ell^H[k]\widetilde{\boldsymbol{C}}_i[k]\boldsymbol{V}_\ell[k], \tag{29}$$

where $\ell\in\{1,2\}$, $k \in \mathcal{K}$, and $i \in \{1,\ldots,M_{\mathrm{R}}\}$. Then, to determine the power allocation matrices, we proceed as follows. If the objective is q_{sum} from (18) (thus, $\ell = 1$), then we solve:

$$\mathbb{P}'_1: \min_{\{\boldsymbol{\Lambda}[k]\in\mathbb{D}^{M_{\mathrm{S}}}\}} \ \mathrm{tr}\sum_{i=1}^{M_{\mathrm{R}}}\sum_{k\in\mathcal{K}}\boldsymbol{\Lambda}[k]\boldsymbol{\Xi}_{i,1}[k] \tag{30a}$$

$$\text{s.t.} \ \ \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{\Lambda}[k] \le K \tag{30b}$$

$$\frac{1}{K}\sum_{k\in\mathcal{K}}\log|\boldsymbol{I}_{M_{\mathrm{S}}}+\gamma\boldsymbol{\Lambda}[k]\boldsymbol{\Omega}_1[k]| \ge t \tag{30c}$$

$$\boldsymbol{\Lambda}[k] \succeq \boldsymbol{0}, \ \forall k \in \mathcal{K} \tag{30d}$$

If the objective is $q_{\max}$ from (19) (thus, $\ell = 2$), we solve:

$$\mathbb{P}'_2: \min_{\{\boldsymbol{\Lambda}[k]\in\mathbb{D}^{M_{\mathrm{S}}}\}} \ \beta' \tag{31a}$$

$$\text{s.t.} \ \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{\Lambda}[k]\boldsymbol{\Xi}_{i,2}[k] \le \beta', \ 1 \le i \le M_{\mathrm{R}}, \tag{31b}$$

$$\frac{1}{K}\sum_{k\in\mathcal{K}}\log|\boldsymbol{I}_{M_{\mathrm{S}}}+\gamma\boldsymbol{\Lambda}[k]\boldsymbol{\Omega}_2[k]| \ge t \tag{31c}$$

$$\text{(30b) and (30d)} \tag{31d}$$

Remark 1 (Convexity)**.** *Since the logdet function is concave in $\{\boldsymbol{\Lambda}[k]\}_{k\in\mathcal{K}}$, the cost functions and the other constraints are convex, and the set of diagonal matrices is convex, both $\mathbb{P}'_1$ and $\mathbb{P}'_2$ are convex.*

Remark 2 (Feasibility)**.** *$\mathbb{P}'_1$ and $\mathbb{P}'_2$ need not be feasible, even though $t = R(M_S)/\rho$ has been set by taking into account the predefined number of streams M_S. Feasibility depends on the existence of power allocation matrices simultaneously satisfying the QoS requirement and the TX power constraint. Specifically, for $\ell = \{1,2\}$, let*

$$\widetilde{\mathcal{I}}_\ell^{max} = \max_{\{\boldsymbol{\Lambda}[k]\in\mathbb{D}^{M_S}\}} \ \frac{1}{K}\sum_{k\in\mathcal{K}}\log|\boldsymbol{I}_{M_S}+\gamma\boldsymbol{\Lambda}[k]\boldsymbol{\Omega}_\ell[k]| \tag{32a}$$

$$\text{s.t.} \ \ \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{\Lambda}[k] \le K \tag{32b}$$

Then, problem $\mathbb{P}'_\ell$ is feasible iff $\widetilde{\mathcal{I}}_\ell^{max} \ge t$, $\ell \in \{1,2\}$.

B. The Single-Stream Case

Although both $\mathbb{P}'_1$ and $\mathbb{P}'_2$ are convex, they do not admit a closed-form solution and have to be solved numerically. To get some insight on the structure of the optimal power allocation, let us focus on $\mathbb{P}'_1$ and the case of a single data stream, *i.e.,* $M_{\mathrm{S}} = 1$, for which the following *water-filling-like* solution emerges.

Lemma 2. *Let $c[k] \triangleq \mathrm{tr}\sum_{i=1}^{M_R}\boldsymbol{V}_1^H[k]\widetilde{\boldsymbol{C}}_i[k]\boldsymbol{V}_1[k] \ge 0$. For $M_S = 1$, note that matrices $\boldsymbol{\Omega}_1[k]$ and $\boldsymbol{\Lambda}[k]$ reduce to scalars $\omega_1[k]$ and $\lambda[k]$, respectively. Assuming $\widetilde{\mathcal{I}}_1^{max} \ge t$ so that the problem is feasible, the optimal power allocation solving $\mathbb{P}'_1$ for $M_S = 1$ is given by*

$$\lambda^{\star}[k] = \left(\frac{\chi}{c[k]+\mu} - \frac{1}{\gamma\omega_1[k]}\right)^+, \ \forall k \in \mathcal{K} \tag{33}$$

where the scalars $\mu \ge 0$ and $\chi \ge 0$ are the Lagrange multipliers associated with the transmit power and MI constraints, respectively. The MI constraint must be tight at the optimal solution, and thus $\chi > 0$; whereas the transmit power constraint may or may not be tight.

Proof. See Appendix C. □

It is seen that (33) is reminiscent of the standard water-filling solution, with the difference that now the "water level" $\chi/(c[k]+\mu)$ is not constant across subcarriers. Note that if the k-th subcarrier SI channel coefficient $c[k]$ is sufficiently large, or the SNR $\gamma\omega_1[k]$ is sufficiently small, then $\lambda^{\star}[k] = 0$ and no power is allocated to that subcarrier.

Inspired by recent views on water-filling algorithms [64], we propose a numerical approach to efficiently find $\{\lambda^{\star}[k]\}_{k\in\mathcal{K}}$, which is given in Algorithm 1 and outlined next. Let $\mathcal{A} \subset \mathcal{K}$ denote the subset of active subcarriers, *i.e.,* the set of indices k for which $\lambda^{\star}[k] > 0$. Substituting (33) into the MI constraint (which must hold with equality), the associated Lagrange multiplier χ is found as

$$\chi = \exp\left\{\frac{Kt}{|\mathcal{A}|} - \frac{1}{|\mathcal{A}|}\sum_{k\in\mathcal{A}}\log\left(\frac{\gamma\omega_1[k]}{c[k]+\mu}\right)\right\}. \tag{34}$$

First, one must check whether the power constraint is inactive. For this, we set $\mu = 0$, initialize $\mathcal{A} = \mathcal{K}$, and iterate between

Algorithm 1: Power Allocation for $M_{\mathrm{S}} = 1$

Input: $\{c[k], w_1[k]\}_{k\in\mathcal{K}}$, t, γ, K, $\mathcal{K}$, $\mu_{\max}$, $\mu_{\min}$, ϵ
Output: $\{\lambda^\star[k]\}_{k\in\mathcal{K}}$

Set $\mathcal{A} \leftarrow \mathcal{K}$ and $\mu \leftarrow 0$;
repeat
  Compute χ via (34) and $\{\lambda^\star[k]\}_{k\in\mathcal{K}}$ via (33);
  Set $\mathcal{A} \leftarrow \{k \,|\, \lambda^\star[k] > 0\}$;
until *convergence*;
if $\sum_{k\in\mathcal{K}} \lambda^\star[k] \le K$ **then**
  Stop;
else
  Set $\mathcal{A} \leftarrow \mathcal{K}$ and $\mu \leftarrow (\mu_{\max} + \mu_{\min})/2$;
  repeat
    Compute χ via (34) and $\{\lambda^\star[k]\}_{k\in\mathcal{K}}$ via (33);
    Set $\mathcal{A} \leftarrow \{k \,|\, \lambda^\star[k] > 0\}$;
  until *convergence*;
  repeat
    if $\sum_{k\in\mathcal{K}} \lambda^\star[k] < K$ **then**
      $\mu_{\max} \leftarrow (\mu_{\max} + \mu_{\min})/2$;
    else
      $\mu_{\min} \leftarrow (\mu_{\min} + \mu_{\min})/2$;
    end
    Set $\mathcal{A} \leftarrow \mathcal{K}$ and $\mu \leftarrow (\mu_{\max} + \mu_{\min})/2$;
    repeat
      Compute χ via (34) and $\{\lambda^\star[k]\}_{k\in\mathcal{K}}$ via (33);
      Set $\mathcal{A} \leftarrow \{k \,|\, \lambda^\star[k] > 0\}$;
    until *convergence*;
  until $\left|\sum_{k\in\mathcal{K}} \lambda^\star[k] - K\right| \le \epsilon$;
end

(33) and (34) to recompute χ and $\mathcal{A}$. Upon convergence, it must be checked whether the obtained values $\{\lambda^\star[k]\}$ meet the power constraint, in which case the procedure is stopped. Otherwise, at the solution the power constraint must be tight, and the corresponding value $\mu > 0$ needs to be found. This is approached by repeating the above iteration (33)-(34) with values of μ obtained via bisection search.

C. Second Approach: Successive Orthogonalization

In scenarios with stringent QoS demands (large t), the strategy in Sec. IV-A may be unfeasible. As an alternative, we propose a suboptimal approach based on successively modifying the Maximum MI precoders by imposing orthogonality to certain directions of the SI channel while simultaneously meeting design constraints (17b)–(17c).

We initialize $\boldsymbol{F}[k] = \boldsymbol{V}[k]\boldsymbol{\Lambda}^{\frac{1}{2}}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{S}}}$, $\forall k \in \mathcal{K}$, as the precoders solving (26) for $s = M_{\mathrm{S}}$. Note that $\{\boldsymbol{\Lambda}[k]\}$ are computed from the M_{S} largest singular values of the intended subchannels $\{\boldsymbol{H}_1[k]\}$ via water-filling with power constraint K. Thus, assuming $t < R(M_{\mathrm{S}})$ so that the problem is feasible, design constraints (17b)–(17c) will be satisfied, and the maximum MI is obtained; however, these precoders do not take SI into account at all.

Now, the idea is to iteratively refine the precoders to reduce the SI objective q_{sum}, while ensuring that the MI constraint is still satisfied. To this end, we impose that, for each $k \in \mathcal{K}$, the new precoder be orthogonal to a subset of eigenvectors (termed the *protected directions*) of the total SI matrix

$$\widetilde{\boldsymbol{C}}[k] = \sum_{i=1}^{M_{\mathrm{R}}} \widetilde{\boldsymbol{C}}_i[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{T}}}, \tag{35}$$

Algorithm 2: Successive Orthogonalization

Input: $\{\widetilde{\boldsymbol{C}}[k]\}_{k\in\mathcal{K}}$, $\{\boldsymbol{H}_1[k]\}_{k\in\mathcal{K}}$, M_{S}, t, K, $\mathcal{K}$
Output: $\{\boldsymbol{F}[k]\}_{k\in\mathcal{K}}$

$\boldsymbol{V}[k] \leftarrow M_{\mathrm{S}}$ dominant RSVs of $\boldsymbol{H}_1[k]$, $\forall k \in \mathcal{K}$;
$\boldsymbol{\Lambda}[k] \leftarrow$ Water-Filling($\{\boldsymbol{H}_1[k]\}_{k\in\mathcal{K}}$, K);
$\boldsymbol{F}[k] \leftarrow \boldsymbol{V}[k]\boldsymbol{\Lambda}^{\frac{1}{2}}[k]$, $\forall k \in \mathcal{K}$, $\mathcal{P}_0 \leftarrow \emptyset$;
Compute SI cost q_0 via (36) and set $q \leftarrow q_0$;
Compute the EVD of $\widetilde{\boldsymbol{C}}[k]$, $\forall k \in \mathcal{K}$, and set $i \leftarrow 1$;
repeat
  Compute $\boldsymbol{H}_{1,i}[k]$ via (37)–(38), $\forall k \in \mathcal{K}$;
  $\boldsymbol{V}[k] \leftarrow M_{\mathrm{S}}$ dominant RSVs of $\boldsymbol{H}_{1,i}[k]$, $\forall k \in \mathcal{K}$;
  $\boldsymbol{\Lambda}[k] \leftarrow$ Water-Filling($\{\boldsymbol{H}_{1,i}[k]\}_{k\in\mathcal{K}}$, K);
  $\widetilde{\boldsymbol{F}}[k] \leftarrow \boldsymbol{V}[k]\boldsymbol{\Lambda}^{\frac{1}{2}}[k]$, $\forall k \in \mathcal{K}$;
  Compute SI cost q_i via (36);
  if (40) *holds* **and** $q_i < q$ **then**
    $\mathcal{P}_i \leftarrow \mathcal{P}_{i-1} \cup \{i\}$;
    $\boldsymbol{F}[k] \leftarrow \widetilde{\boldsymbol{F}}[k]$, $\forall k \in \mathcal{K}$;
    $q \leftarrow q_i$;
  end
  $i \leftarrow i + 1$;
until $i = M_{\mathrm{T}}$ or $|\mathcal{P}_i| = M_{\mathrm{T}} - M_{\mathrm{S}}$;

since (18) can be rewritten as

$$q_{\mathrm{sum}} = \mathrm{tr} \sum_{k\in\mathcal{K}} \boldsymbol{F}^H[k]\widetilde{\boldsymbol{C}}[k]\boldsymbol{F}[k]. \tag{36}$$

While this strategy has the potential to decrease SI, it also reduces the achievable MI, and could result in violation of the MI constraint. Letting $\mathcal{P}[k] \subset \{1, \ldots, M_{\mathrm{T}}\}$ be the index set of protected directions for the k-th subcarrier, finding the optimal configuration $\{\mathcal{P}[k]\}$ that minimizes SI while meeting the MI and TX power constraints becomes a combinatorial problem. To sidestep this issue, we propose to (i) have common protected directions for all subcarriers, *i.e.,* $\mathcal{P}[k] = \mathcal{P}$ $\forall k \in \mathcal{K}$; (ii) construct the set $\mathcal{P}$ using a greedy search with the procedure sketched in Algorithm 2 and described next.

Let $\boldsymbol{v}_j[k]$ be the unit-norm eigenvector of $\widetilde{\boldsymbol{C}}[k]$ associated to its j-th largest eigenvalue. We initialize $\mathcal{P}_0$ as the empty set, and for $i = 1, \ldots, M_{\mathrm{T}}$, we construct the following orthogonal projection matrices for all $k \in \mathcal{K}$:

$$\boldsymbol{P}_i^{\perp}[k] = \boldsymbol{I}_{M_{\mathrm{T}}} - \left(\sum_{j\in\mathcal{P}_{i-1}} \boldsymbol{v}_j[k]\boldsymbol{v}_j^H[k] + \boldsymbol{v}_i[k]\boldsymbol{v}_i^H[k] \right), \tag{37}$$

and, from them, the *effective* intended channel matrices

$$\boldsymbol{H}_{1,i}[k] = \boldsymbol{H}_1[k]\boldsymbol{P}_i^{\perp}[k], \ \forall k \in \mathcal{K}. \tag{38}$$

The (tentative) precoders $\widetilde{\boldsymbol{F}}[k] = \boldsymbol{V}[k]\boldsymbol{\Lambda}^{\frac{1}{2}}[k]$ at iteration i are then obtained by solving

$$\max_{\{\widetilde{\boldsymbol{F}}[k]\}} \frac{1}{K} \sum_{k\in\mathcal{K}} \log \left| \boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma \boldsymbol{H}_{1,i}[k]\widetilde{\boldsymbol{F}}[k]\widetilde{\boldsymbol{F}}^H[k]\boldsymbol{H}_{1,i}^H[k] \right| \tag{39a}$$

$$\text{s.t.} \sum_{k\in\mathcal{K}} \|\widetilde{\boldsymbol{F}}[k]\|_F^2 \le K. \tag{39b}$$

In this way, the columns of $\boldsymbol{V}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{S}}}$ are taken as the M_{S} dominant right singular vectors (RSVs) of $\boldsymbol{H}_{1,i}[k]$, whereas $\boldsymbol{\Lambda}[k] \in \mathbb{C}^{M_{\mathrm{S}}\times M_{\mathrm{S}}}$ is found via water-filling. Note that, by construction, $\boldsymbol{H}_{1,i}[k]\boldsymbol{v}_j[k] = \boldsymbol{0}$ for $j \in \mathcal{P}_{i-1} \cup \{i\}$. This implies that all RSVs of $\boldsymbol{H}_{1,i}[k]$, and in particular the columns of $\widetilde{\boldsymbol{F}}[k]$, are orthogonal to $\boldsymbol{v}_j[k]$ for $j \in \mathcal{P}_{i-1} \cup \{i\}$, as desired.

At this point it is necessary to verify whether the tentative precoders actually meet the MI constraint, by checking whether the attained maximum in (39a), say $\widetilde{\mathcal{I}}_i$, satisfies

$$\widetilde{\mathcal{I}}_i \geq t. \tag{40}$$

In that case, we check if the total SI has also been reduced with respect to the previous iteration, and if so, we update $\mathcal{P}_i = \mathcal{P}_{i-1} \cup \{i\}$; otherwise, $\mathcal{P}_i = \mathcal{P}_{i-1}$.

Remark 3 (Stopping criterion)**.** *Algorithm 2 iterates until either $i = M_T$ or exactly $M_T - M_S$ protected directions have been found, whichever comes first. The latter condition guarantees that the preserved spatial DoF can support the transmission of M_S data streams per subcarrier.*

Remark 4 (Worst-case performance)**.** *If the set of protected directions remains empty ($\mathcal{P} = \emptyset$) after M_T iterations, then no orthogonalization direction can reduce the total SI* (18) *without violating* (17b)–(17c) *for the prescribed number of data streams per subcarrier M_S. In such case, Algorithm 2 provides the conventional maximum MI configuration, ensuring that the total SI is upper bounded by that of the maximum MI baseline.*

In fact, Algorithm 2 bridges between the Maximum MI solution and conventional Spatial Nulling (SN) [23], which focuses on maximum total SI reduction without regarding the achieved MI. Note from (14) that, in the absence of transmit noise (*i.e.,* $\eta_{\mathrm{T}} \to \infty$), $\widetilde{\boldsymbol{C}}_i[k]$ is a rank-1 matrix, so that $\widetilde{\boldsymbol{C}}[k]$ in (35) becomes rank deficient (with rank M_{R}) if $M_{\mathrm{T}} > M_{\mathrm{R}}$. Then, by constraining the precoder $\boldsymbol{F}[k]$ to lie in the null space of $\widetilde{\boldsymbol{C}}[k]$, SI would be completely avoided; this is the basis of conventional SN. With transmit noise, $\widetilde{\boldsymbol{C}}[k]$ is full rank in general, and the SN approach can be generalized by constraining $\boldsymbol{F}[k]$ to lie in the span of the $M_{\mathrm{T}} - M_{\mathrm{R}}$ least dominant eigenvectors of $\widetilde{\boldsymbol{C}}[k]$. We still refer to this design as SN, although the SI is not truly nulled out anymore. In settings with a relaxed MI requirement, the design in Algorithm 2 exhibits similar SIC performance to SN. However, in contrast with SN, the proposed approach (i) does not require that $M_{\mathrm{T}} > M_{\mathrm{R}}$, and (ii) it always guarantees the QoS requirement.

V. Simulation Results

The performance of the proposed designs is evaluated in an upper mid-band [65] setting with a bandwidth $W = 50$ MHz operating at carrier frequency $f_c = 10$ GHz. The number of active subcarriers is $K = 100$ (subcarrier spacing $\Delta_f = \frac{W}{K} = 500$ kHz), with CP redundancy of 25%. We assume Uniform Linear Arrays (ULAs) with $\frac{\lambda}{2}$-spaced elements (λ is the carrier wavelength). The FD node (BS) geometry is as in [66, Fig. 2] with $\alpha = \beta = \frac{\pi}{2}$ and $\delta = 100\lambda$. Unless otherwise stated, we consider $M_{\mathrm{T}} = 16$, $M_{\mathrm{R}} = M_{\mathrm{R}'} = 8$, and $\eta_{\mathrm{T}} = 40$ dB.

The intended channel $\boldsymbol{H}_1[k]$ follows a frequency-selective Saleh-Valenzuela model, as in (41), with $N_{\mathrm{c}} = 7$ clusters and $N_{\mathrm{r}} = 3$ rays per cluster. In (41), α_{mn}, τ_{mn}, ϑ_{mn}, and θ_{mn} denote the complex path gain, path delay, direction of departure (DoD), and direction of arrival (DoA), respectively, whereas $\boldsymbol{a}_{\mathrm{T}}$ and $\boldsymbol{a}_{\mathrm{R}}$ stand for the TX and RX ULA steering vectors. A maximum delay $\tau_{\max} = 500$ ns is assumed, so that $D = \lceil W\tau_{\max} \rceil = 25$. Filtering effects in (41) are modeled with a 50%-excess bandwidth raised cosine filter $h_{\mathrm{I}}(t)$. In the simulations, DoD/DoA are Gaussian distributed with a standard deviation of 23° and mean cluster angles uniformly distributed within $[0, 360^\circ]$. Path gains are i.i.d. zero-mean unit-variance complex Gaussian distributed, and delays are uniformly distributed within $[0, D]$. The intended channel matrices are normalized such that $\sum_{k\in\mathcal{K}} \|\boldsymbol{H}_1[k]\|_{\mathrm{F}}^2 = M_{\mathrm{T}} M_{\mathrm{R}'}$.

Since the transmit and receive arrays of the FD node are co-located, we consider a Rice model for the SI channel, where the Line-of-Sight (LOS) component models the effect of the spherical wavefront in the near field, as in [41], [67]:

$$\left[\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LOS}}\right]_{ij} = \frac{c}{r_{ij}} \exp\left\{-\imath 2\pi \frac{r_{ij}}{\lambda}\right\}, \tag{42}$$

with r_{ij} the distance between the i-th TX and j-th RX antennas, and c a normalization constant to have $\|\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LOS}}\|_{\mathrm{F}}^2 = \frac{M_{\mathrm{T}} M_{\mathrm{R}}}{K}$. The Non-Line-of-Sight (NLOS) component $\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}[k]$ captures the effect of nearby reflectors, and follows the frequency-selective multipath model (41) with $N_{\mathrm{c}} = 7$, $N_{\mathrm{r}} = 3$. $\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}[k]$ are normalized such that $\sum_{k\in\mathcal{K}} \|\boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}[k]\|_{\mathrm{F}}^2 = M_{\mathrm{T}} M_{\mathrm{R}}$. Then, the SI channel matrices $\boldsymbol{H}_{\mathrm{SI}}[k]$ are obtained as

$$\boldsymbol{H}_{\mathrm{SI}}[k] = \sqrt{\frac{\kappa}{\kappa+1}} \boldsymbol{H}_{\mathrm{SI}}^{\mathrm{LOS}} + \sqrt{\frac{1}{\kappa+1}} \boldsymbol{H}_{\mathrm{SI}}^{\mathrm{NLOS}}[k], \tag{43}$$

which are normalized so that $\sum_{k\in\mathcal{K}} \|\boldsymbol{H}_{\mathrm{SI}}[k]\|_{\mathrm{F}}^2 = M_{\mathrm{T}} M_{\mathrm{R}}$. In the sequel, we will consider a Rice factor $\kappa = 10$ dB.

Recalling (1) and (26), we introduce the *normalized performance loss* (NPL) corresponding to the MI requirement:

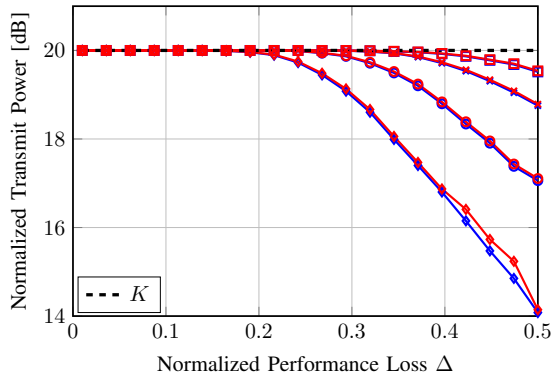
$$\Delta \triangleq \frac{R(d) - t}{R(d)} = 1 - \frac{1}{\rho} \in [0, 1). \tag{44}$$

All convex optimization problems have been implemented in MATLAB using CVX toolbox [68].

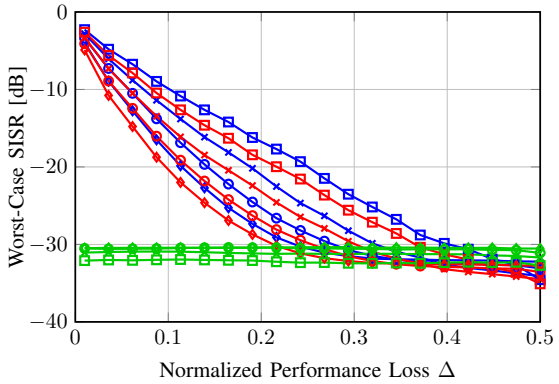
A. Results for SDR-based Designs

Fig. 2 illustrates the performance of the SDR-based schemes from Sec. III, in which not only the precoding matrices but also the number of data streams per subcarrier are optimized. We shall refer to the SDR-based designs having as objective the total SI q_{sum} and worst-case SI $q_{\max}$ as "$\mathbb{P}_1$" and "$\mathbb{P}_2$" respectively, in reference to the corresponding optimization problems from Sec. III. First, we assess the tightness of the power constraint. As seen in Fig. 2a, for small NPL values it becomes necessary for the FD node to transmit at full power to meet the MI constraint. As this constraint is relaxed, *i.e.,* as Δ is increased, it becomes possible to reduce the TX power, and

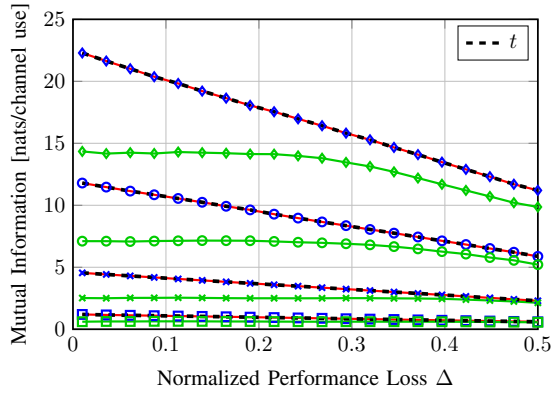
$$\boldsymbol{H}[k] = \sum_{d=0}^{D-1} \sum_{m=0}^{N_{\mathrm{c}}-1} \sum_{n=0}^{N_{\mathrm{r}}-1} \alpha_{mn} h_{\mathrm{I}}(dT_{\mathrm{s}} - \tau_{mn}) \boldsymbol{a}_{\mathrm{R}}(\theta_{mn}) \boldsymbol{a}_{\mathrm{T}}^H(\vartheta_{mn}) e^{-\imath \frac{2\pi k d}{N}} \tag{41}$$



(a) Normalized Transmit Power $\sum_{k\in\mathcal{K}} \|\boldsymbol{F}[k]\|_F^2$.



(b) Worst-Case SI Suppression Ratio (SISR) (47).



(c) Achievable Mutual Information (23). For clarity, results for problems $\mathbb{P}_1$ and $\mathbb{P}_1$ appear as blue markers and red solid lines, respectively.

Fig. 2: Performance assessment of the solutions to $\mathbb{P}_1$ (blue) and $\mathbb{P}_2$ (red), compared with SN (green), at $\gamma = 5$ dB (squares), $\gamma = 15$ dB (crosses), $\gamma = 25$ dB (circles), and $\gamma = 35$ dB (diamonds).

hence the SI, while still satisfying the MI requirement, more so for larger values of the SNR γ at the intended receiver.

As benchmarks, we have considered the Maximum MI (MaxMI) and Spatial Nulling (SN) designs, which can be summarized as follows. Let $\boldsymbol{G}[k] \in \mathbb{C}^{M_{\mathrm{R}'} \times M_{\mathrm{T}}}$ be the adopted intended channel matrix depending on the particular design:

$$\boldsymbol{G}[k] = \begin{cases} \boldsymbol{H}_1[k] & \text{for MaxMI,} \\ \boldsymbol{H}_1[k]\left(\boldsymbol{I}_{M_{\mathrm{T}}} - \widetilde{\boldsymbol{V}}_C[k]\widetilde{\boldsymbol{V}}_C^H[k]\right) & \text{for SN,} \end{cases} \tag{45}$$

with $\widetilde{\boldsymbol{V}}_C[k] \in \mathbb{C}^{M_{\mathrm{T}} \times r}$ containing the $r = \min(M_{\mathrm{T}}, M_{\mathrm{R}})$ dominant eigenvectors of $\widetilde{\boldsymbol{C}}[k]$ in (35). Then, the two considered benchmarks are the solution to

$$\max_{\{\boldsymbol{F}[k]\}} \ \frac{1}{K}\sum_{k\in\mathcal{K}} \log\left|\boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma \boldsymbol{G}[k]\boldsymbol{F}[k]\boldsymbol{F}^H[k]\boldsymbol{G}^H[k]\right| \tag{46a}$$

$$\text{s.t.} \ \ \mathrm{tr}\sum_{k\in\mathcal{K}} \boldsymbol{F}[k]\boldsymbol{F}^H[k] \le \overline{P}_{\mathrm{cons}}(\Delta) \tag{46b}$$

where, for a fair comparison, the TX power constraint $\overline{P}_{\mathrm{cons}}(\Delta)$ in (46b) is set to the corresponding TX power achieved with the proposed SDR-based designs, according to Fig. 2a; and the number of transmitted data streams per subcarrier, *i.e,* the number of columns of $\boldsymbol{F}[k]$, is fixed to the rank of the SDR-based solution to $\mathbb{P}_1$.

To evaluate the SIC capability of the different designs, we adopt the worst-case SI Suppression Ratio (SISR), as in [30]. Letting $P_{\mathrm{SI},i,\mathrm{SIC}}$ and $P_{\mathrm{SI},i,\text{no-SIC}}$ be the SI power at the i-th receive antenna of the FD node with and without SIC capabilities, respectively, the worst case SISR is defined as

$$\mathrm{SISR}_{\mathrm{worst}} \triangleq \frac{\max_{i\in\{1,\ldots,M_{\mathrm{R}}\}} P_{\mathrm{SI},i,\mathrm{SIC}}}{\max_{i\in\{1,\ldots,M_{\mathrm{R}}\}} P_{\mathrm{SI},i,\text{no-SIC}}}, \tag{47}$$

where the $P_{\mathrm{SI},i,\text{no-SIC}}$ is computed as in (19) adopting the MaxMI precoders.

Fig. 2b shows the $\mathrm{SISR}_{\mathrm{worst}}$ achieved by the different designs proposed in this work and by SN. As expected, the worst-case-SI-oriented design $\mathbb{P}_2$ outperforms the total-SI-oriented design $\mathbb{P}_1$ in this regard, although the improvement is slight. This, together with the computational advantage for $M_{\mathrm{R}'} < M_{\mathrm{T}}$ described in Sec. III-D, makes the total-SI-oriented design $\mathbb{P}_1$ particularly attractive. Regarding SN, note that better SIC performance is achieved for smaller values of γ. According to (46), the SN solution is obtained via water-filling over equivalent channels $\boldsymbol{G}[k]$, *i.e.*, the channels projected onto the least significant modes of $\widetilde{\boldsymbol{C}}[k]$. At low SNR (small γ), the water-filling solution tends to concentrate the transmit power into a few streams and subcarriers, which leads to a reduced $\mathrm{SISR}_{\mathrm{worst}}$. In contrast, as γ increases, the water-filling solution activates more streams and subcarriers, thus increasing the level of induced SI. This behavior contrasts with that of the proposed approaches: the solutions to $\mathbb{P}_1$-$\mathbb{P}_2$ tend to reduce the transmit power when the channel conditions are favorable enough to easily satisfy the MI requirement (high γ).

The proposed approaches bridge between the MaxMI solution (*i.e.*, $\mathrm{SISR}_{\mathrm{worst}} = 0$ dB), which ignores SI, and the SN design, which attempts to cancel as much SI as possible, achieving a reduction of about 30 dB with respect to MaxMI; for large values of Δ, the proposed schemes even slightly outperform SN in terms of SIC. The other aspect of the design, namely the achieved MI, is shown in Fig. 2c, where it is seen that the good SIC performance of SN is at the cost of significant MI degradation: for $\gamma = 25$ dB and 35 dB, SN only achieves about 48% and 66% of the MI values attained by MaxMI, respectively. In contrast, the proposed designs provide a much better tradeoff. For instance, for $\Delta = 0.2$ and $\gamma = 35$ dB, $\mathbb{P}_1$-$\mathbb{P}_2$ provide an MI improvement of about 30% with respect to SN, with a degradation of only 2 dB in terms of worst-case SI; for sufficiently large values of Δ, they even outperform SN in terms of both SIC and MI. We observed

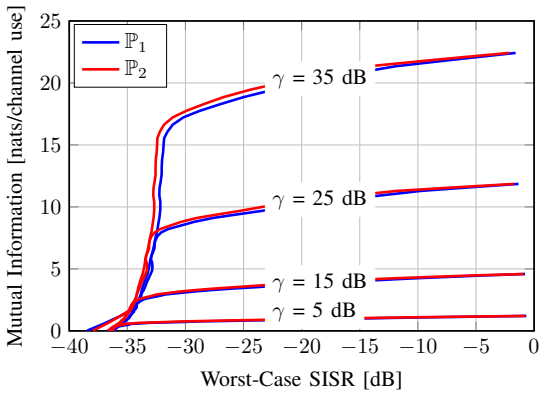


Fig. 3: Worst-case SISR–Achievable MI tradeoff curves.

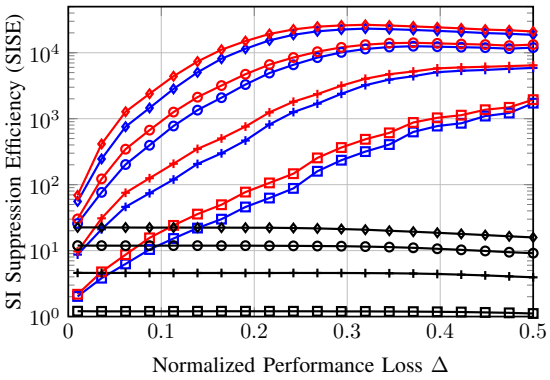


Fig. 4: Self-Interference Suppression Efficiency (SISE) achieved by the optimal solutions to $\mathbb{P}_1$ (blue) and $\mathbb{P}_2$ (red), and MaxMI (black) precoders, at $\gamma = 5$ dB (squares), $\gamma = 15$ dB (crosses), $\gamma = 25$ dB (circles), and $\gamma = 35$ dB (diamonds).

that the optimal number of streams produced by the SDR-based designs roughly fluctuated between 2 and 3 when $\gamma = 5$ dB, between 3 and 4 when $\gamma = 15$ dB, between 4 and 5 when $\gamma = 25$ dB, and between 5 and 6 when $\gamma = 35$ dB.

Fig. 3 further illustrates the tradeoff between achievable MI and worst-case SISR. As the required SIC level is increased, meaning that a smaller worst-case SISR is achieved, the ensuing MI degradation remains quite graceful, until reaching a regime (left side of Fig. 3) in which further SIC can only be attained by reducing the transmit power, resulting in a sharp drop in MI. Again, no significant difference is observed between the performances of the $\mathbb{P}_1$ and $\mathbb{P}_2$ designs.

For a more clear interpretation of the results, we define the Self-Interference Suppression Efficiency (SISE) as

$$\text{SISE} \triangleq \frac{\mathcal{I}}{\text{SISR}_{\text{worst}}} \mathbb{1}_{\mathcal{I} \geq t}, \tag{48}$$

i.e., the ratio of the achieved MI to the $\text{SISR}_{\text{worst}}$, weighted by the indicator function $\mathbb{1}_{\mathcal{I} \geq t}$, which equals one if $\mathcal{I} \geq t$ and zero otherwise. This efficiency metric measures the amount of information transmitted per unit of SI suppression, thus quantifying the fundamental MI-SI tradeoff. Note that only those precoding techniques satisfying the minimum MI requirement have nonzero efficiency, since otherwise they would not be feasible solutions to the problem at hand. The SISE is evaluated in Fig. 4 in terms of the NPL Δ. Observe that only the proposed SDR-based designs and MaxMI are shown since SN violates the MI constraint for all values

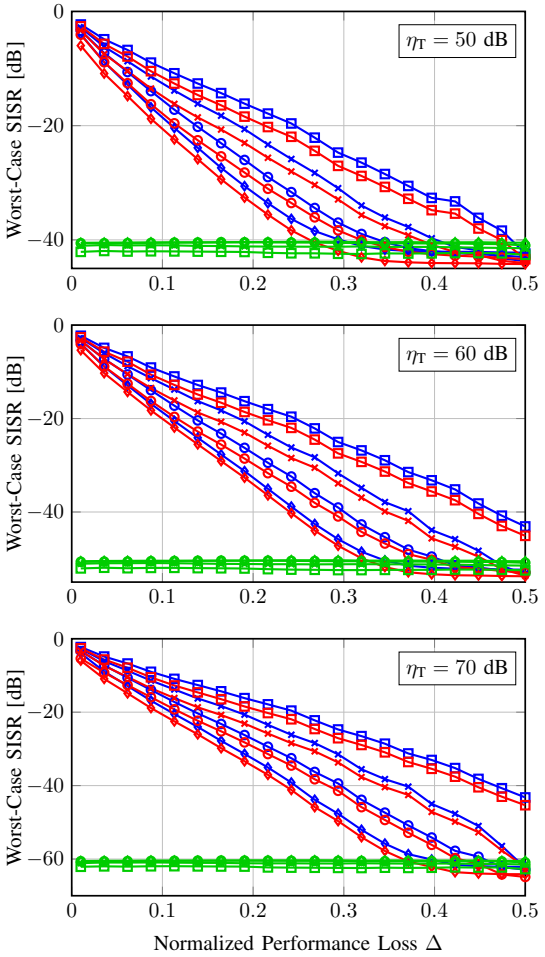


Fig. 5: Worst-case SISR performance of $\mathbb{P}_1$ (blue) and $\mathbb{P}_2$ (red), and SN (green), at $\gamma = 5$ dB (squares), $\gamma = 15$ dB (crosses), $\gamma = 25$ dB (circles), and $\gamma = 35$ dB (diamonds), for different values of the TX SNR η_{T}.

of Δ, and thus SISE $= 0$. As expected, since MaxMI has no SI suppression capability (*i.e.*, $\text{SISR}_{\text{worst}} = 0$ dB), it underperforms both proposed SDR-based designs $\mathbb{P}_1$ and $\mathbb{P}_2$ in terms of SISE. This efficiency metric corroborates that both $\mathbb{P}_1$ and $\mathbb{P}_2$ achieve a comparable MI-SI tradeoff. Both solutions satisfy the MI constraint with strict equality (see Fig. 2c), while exhibiting similar SI suppression capabilities (see Fig. 2b). This close alignment in SISE is of high practical value: when computational complexity is the limiting factor, the less costly approach described in Sec. III-D based on Corollary 1 provides an efficient solution balancing the MI-SI tradeoff without sacrificing performance. For instance, in the considered setting, solving problem $\mathbb{P}_1$ via Corollary 1 reduces the computation time in a standard laptop by about 40% compared to solving it directly. This runtime saving further increases as the ratio $M_{\text{T}}/M_{\text{R}'}$ and/or K grow. Since specific runtimes depend on the underlying hardware architecture, reduced execution times can be expected when using commercial high-performance signal processing platforms.

The impact of η_{T}, the SNR at each TX antenna, is assessed in Fig. 5. As η_{T} increases, the impact of TX noise becomes less pronounced, see (8), and all methods achieve a better $\text{SISR}_{\text{worst}}$ for a given Δ (with respect to Fig. 2b, corresponding to $\eta_{\text{T}} = 40$ dB); also, both $\mathbb{P}_1$-$\mathbb{P}_2$ require a larger Δ to achieve the SIC

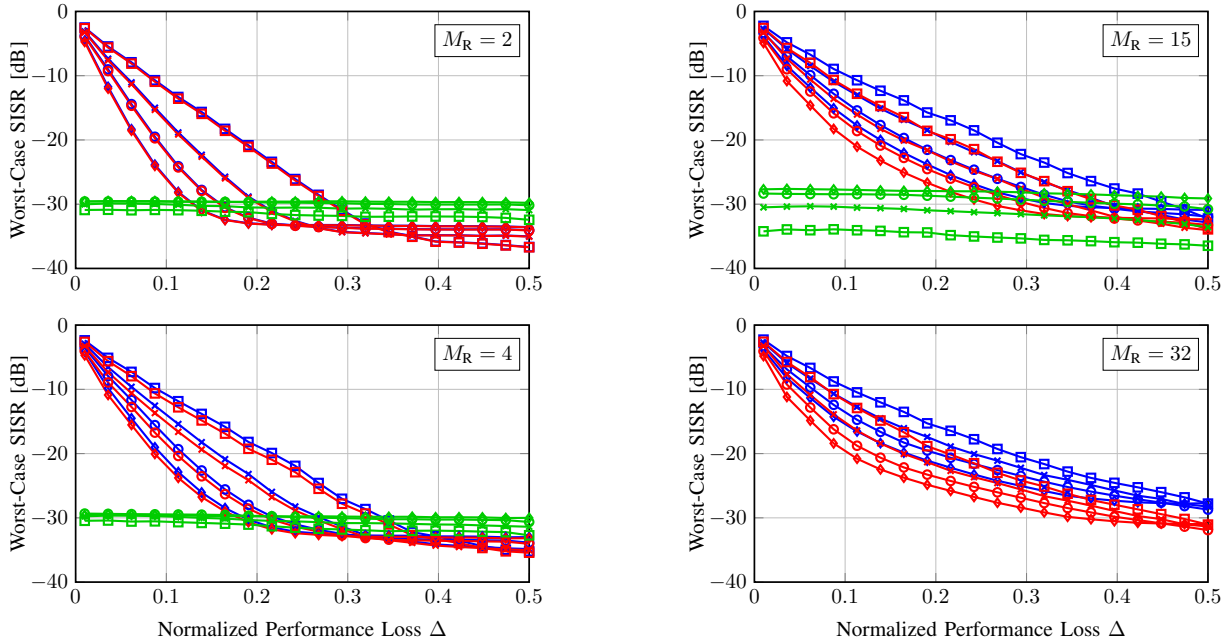


Fig. 6: Worst-case SISR performance of $\mathbb{P}_1$ (blue) and $\mathbb{P}_2$ (red), and SN (green), at $\gamma = 5$ dB (squares), $\gamma = 15$ dB (crosses), $\gamma = 25$ dB (circles), and $\gamma = 35$ dB (diamonds), for different number of RX antennas M_{R}.

performance of SN. Fig. 6 shows the impact of the RX array size at the FD node, M_{R}, which determines the number of least significant modes of $\widetilde{\boldsymbol{C}}_i[k]$. Observe that, as M_{R} grows, both $\mathbb{P}_1$-$\mathbb{P}_2$ attain the SIC performance of SN at larger Δ, and that $\mathbb{P}_2$ exhibits better SIC performance than $\mathbb{P}_1$. The latter is due to the adopted cost function in each design. Since $\mathbb{P}_1$ minimizes the total SI, its optimal solution can opt to tolerate higher SI on specific antennas if it facilitates satisfying the MI constraint while minimizing the cost function. In contrast, $\mathbb{P}_2$ minimizes the worst-case per-antenna SI, which ensures that the peak SI across all receive antennas is minimized. However, for small M_{R}, $\mathbb{P}_1$ has fewer spatial DoF to compensate for higher SI on specific antennas, and it tends to produce solutions exhibiting similar SIC performance to $\mathbb{P}_2$. As a final comment on Fig. 6, note that the SN design requires $M_{\mathrm{R}} < M_{\mathrm{T}}$ (otherwise, (45) yields $\boldsymbol{G}[k] = \boldsymbol{0}$); remarkably, both $\mathbb{P}_1$-$\mathbb{P}_2$ still produce a viable solution even if $M_{\mathrm{R}} \geq M_{\mathrm{T}}$.

B. Joint vs. Per-Subcarrier Designs

Whereas we have pursued a *joint* approach to precoder design (in the sense that the objective and constraints consider the contribution of all active subcarriers to SI, TX power, and MI), the authors of [39] suggest that their narrowband design can be generalized to wideband scenarios by individually applying it to each subcarrier. For a fair comparison with such *per-subcarrier* approach, we modify the formulation in [39] by replacing its QoS constraint on the power at the intended RX input by one on MI, as follows. Let $q[k]$ be some SI-related function of $\boldsymbol{F}[k]$; then, for each $k \in \mathcal{K}$, we solve

$$\min_{\boldsymbol{F}[k]} \quad q[k] \tag{49a}$$

$$\text{s.t.} \quad \mathrm{tr}\, \boldsymbol{F}[k]\boldsymbol{F}^H[k] \leq 1 \tag{49b}$$

$$\log \left| \boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma \boldsymbol{H}_1[k]\boldsymbol{F}[k]\boldsymbol{F}^H[k]\boldsymbol{H}_1^H[k] \right| \geq t \tag{49c}$$

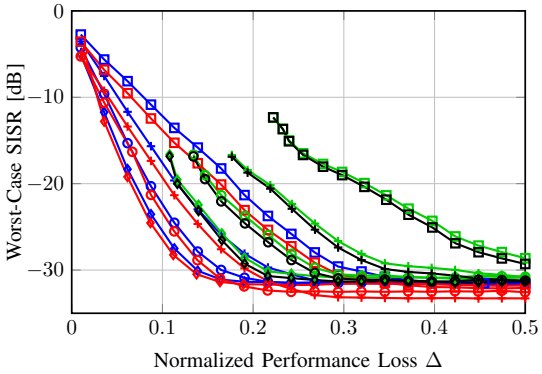


Fig. 7: $\mathrm{SISR}_{\mathrm{worst}}$ achieved by the Joint ($\mathbb{P}_1$ in blue and $\mathbb{P}_2$ in red) and Per-Subcarrier ($\mathbb{P}_1$ in green and $\mathbb{P}_2$ in black) approaches to $\mathbb{P}_1$ and $\mathbb{P}_2$ at $\gamma = 5$ dB (squares), $\gamma = 15$ dB (crosses), $\gamma = 25$ dB (circles), and $\gamma = 35$ dB (diamonds)

The cost functions $q[k]$ are adapted from (18)–(19) as

$$q_{\mathrm{sum}}[k] \quad = \quad \mathrm{tr} \sum_{i=1}^{M_{\mathrm{R}}} \boldsymbol{F}^H[k]\widetilde{\boldsymbol{C}}_i[k]\boldsymbol{F}[k], \tag{50}$$

$$q_{\mathrm{max}}[k] \quad = \quad \max_{i \in \{1,\ldots,M_{\mathrm{R}}\}} \mathrm{tr}\, \boldsymbol{F}^H[k]\widetilde{\boldsymbol{C}}_i[k]\boldsymbol{F}[k]. \tag{51}$$

Results are shown in Fig. 7 in terms of worst-case SI Suppression Ratio. The setting is the same as that of Fig. 2 except that we set $M_{\mathrm{R}'} = 1$ as in [39]. Note that, for a fair comparison, the per-subcarrier solution is declared unfeasible if the aggregate MI does not satisfy the global MI constraint. This infeasibility occurs because the per-subcarrier approach evaluates each subcarrier independently, being incapable of compensating for subcarriers experiencing poor channel conditions or deep fades by exploiting those experiencing better channel conditions. It is clear that, for a given NPL level, the proposed joint designs are much more effective at reducing SI than the corresponding per-subcarrier approaches. Moreover, in terms of computation

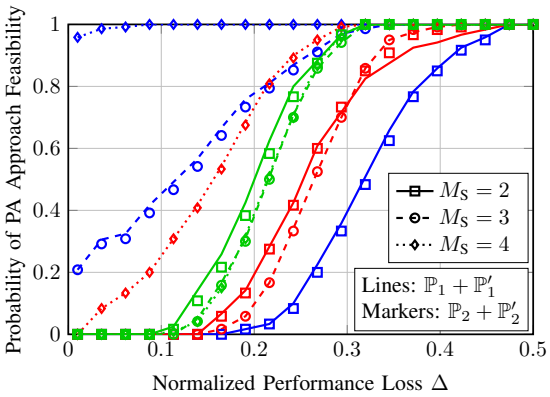


Fig. 8: Empirical probability that the suboptimal approach based on power allocation (Sec. IV-A) is feasible at $\gamma = 15$ dB (blue), $\gamma = 25$ dB (red), and $\gamma = 35$ dB (green).

time, we observed that the proposed designs are about 8 times faster than the per-subcarrier approaches, as the solver can further exploit hidden dependencies between $\{\boldsymbol{F}[k]\}_{k\in\mathcal{K}}$.

C. Fixing the Number of Streams

A feature of the SDR-based designs from Sec. III is that the number of data streams at each subcarrier is optimized together with the precoding matrices. In certain cases, there may be a limit to the number of streams that the transceiver can handle, which motivated the development of the techniques from Sec. IV for a fixed number M_S of streams per subcarrier.

In the first approach, described in Sec. IV-A, the dominant M_S eigenvectors of the SDR-based solutions are kept, and then the power allocation (PA) terms are optimized. As discussed in Sec. IV-A, this approach need not always be feasible. Fig. 8 shows the probability of obtaining a feasible solution in the setting of Fig. 2, in terms of the NPL, redefined now as

$$\Delta \triangleq \frac{R(M_S) - t}{R(M_S)}. \tag{52}$$

It is seen that, unless Δ is sufficiently large, for a given SNR feasibility becomes less likely as the number of streams M_S is reduced with respect to the rank of the optimal solutions to $\mathbb{P}_1$-$\mathbb{P}_2$. This is true for both approaches, $\mathbb{P}'_1$ and $\mathbb{P}'_2$. Whereas the optimal solutions to $\mathbb{P}_1$-$\mathbb{P}_2$ generally yield higher ranks, increasing M_S does not necessarily improve feasibility. Since the MI threshold t scales linearly with M_S but logarithmically with γ, stringent values of Δ can compromise feasibility as γ grows if M_S is kept much lower than the optimal ranks.

This potential lack of feasibility motivated the development of the successive orthogonalization approach discussed in Sec. IV-C. In Fig. 9, we compare worst-case SISR performance of the SDR-based approaches (termed "$\mathbb{P}_\ell$"), their power allocation-based suboptimal counterparts (termed "$\mathbb{P}_\ell + \mathbb{P}'_\ell$"), and the successive orthogonalization method from Algorithm 2. We considered two SNR values: $\gamma = 25$ dB, for which we fixed the number of streams of the suboptimal designs to $M_S = 3$, and $\gamma = 35$ dB, for which $M_S = 4$. As mentioned above, the optimal SDR-based solutions produced higher values of the number of streams in both cases.

Note that, whenever the power allocation-based design from Sec. IV-A is feasible, its SIC performance remains close to that of the corresponding SDR-based optimal approach. Unfortunately, under stringent QoS requirements (small Δ), feasibility is lost. In contrast, the successive orthogonalization scheme from Sec. IV-C is always feasible, providing SIC performance close to the M_S-constrained optimal solutions (*i.e.*, $\mathbb{P}_\ell + \mathbb{P}'_\ell$) for sufficiently large NPL; however, as Δ is decreased, a gap with respect to the optimal solutions appears, which becomes narrower under higher SNR conditions. Also, with larger SNR, the power allocation-based design becomes feasible for a wider range of NPL values.

Thus, Fig. 9 reveals a threefold tradeoff involving feasibility, SI suppression, and complexity. Whereas successive orthogonalization is suboptimal, it always provides a feasible solution that allows to fix the number of data streams per subcarrier, and with reduced computational complexity, as it does not require the use of IPMs. In contrast, the PA-based approach, when feasible, performs close to the optimal solutions; but it requires solving two optimization problems through IPMs, and for feasibility one must have either a sufficiently large SNR or a conservative QoS requirement.

To close this section, we consider the single-stream case $M_S = 1$ to verify the suitability of Algorithm 1 to find the power allocation coefficients of the suboptimal approach from Sec. IV-A. Fig. 10 shows the worst-case SISR achieved by the optimal SDR-based design ($\mathbb{P}_1$) and the suboptimal approaches from Sec. IV: solving $\mathbb{P}'_1$ through an IPM, solving $\mathbb{P}'_1$ through Algorithm 1, and the sucessive orthogonalization approach (Algorithm 2). Clearly, Algorithm 1 provides the same solution as the interior point method, and therefore it constitutes a much better choice from the point of view of computational complexity, especially in settings with a large number of subcarriers. Similarly to our findings in Fig. 9, the power allocation-based approach provides SIC levels close to those of the optimal SDR-based design $\mathbb{P}_1$. Incidentally, note that the differences in the SI curves for $\mathbb{P}_1$ in Fig. 10 and the left columnn of Fig. 9 is due to the fact that, for a given NPL value Δ, the MI requirement t depends on the number of streams considered as $t = (1-\Delta)R(M_S)$; see (52). Since $R(M_S)$ decreases when M_S is reduced, a less stringent MI requirement in absolute terms results, so that the optimal SDR-based design is able to improve its SIC performance.

VI. Conclusions

We have addressed the design of digital SIC precoders for MIMO-OFDM STAR systems. The expressions of the SI power affecting each RX antenna have been developed, accounting for limited TX dynamic range. Two approaches based on the minimization of the total SI and the worst-case SI at the RX array, under a QoS requirement in terms of MI, have been discussed. The non-convexity of the resulting problems was addressed via semidefinite relaxation, which was shown to be tight, providing the optimal precoders and number of data streams per subcarrier. These designs significantly outperform previous per-subcarrier based approaches. For those cases where the transceivers cannot support the optimal number of data streams, two suboptimal designs based on power allocation and successive orthogonalization have been

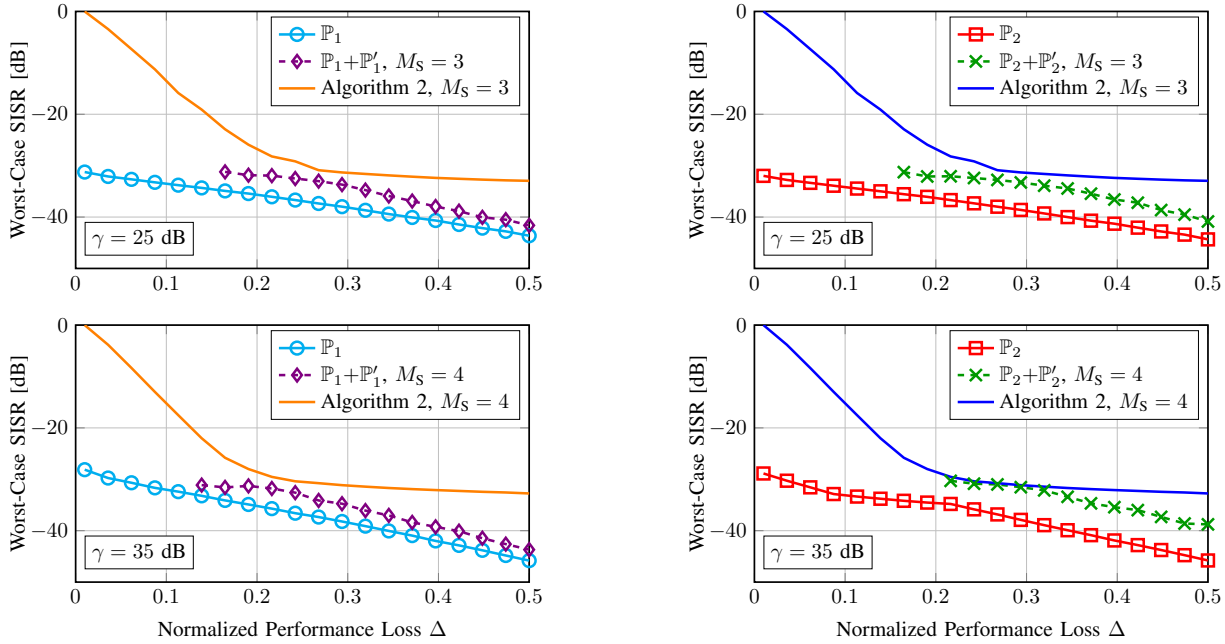


Fig. 9: Performance comparison of the optimal design from Sec. III and the suboptimal approaches based on power allocation (Sec. IV-A) and successive orthogonalization (Sec. IV-C), adopting the total SI criterion (18) (left), and the worst-case SI criterion (19) (right).

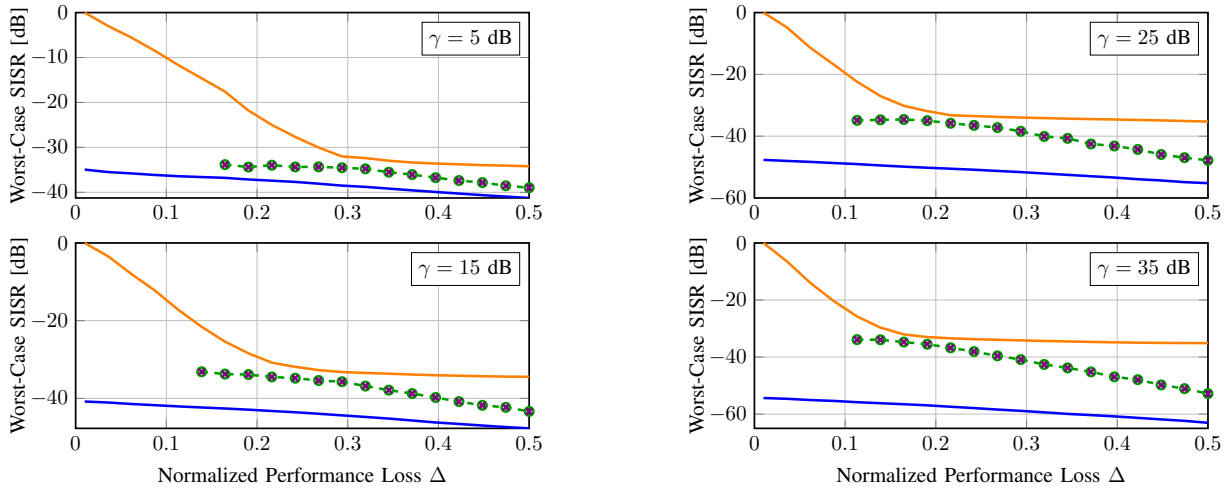


Fig. 10: Performance comparison of the optimal design from Sec. III (blue), the suboptimal approaches based on power allocation from Sec. IV-A (solving $\mathbb{P}'_1$ via an IPM in green, and via Algorithm 1 in purple crosses), and successive orthogonalization from Sec. IV-C (orange), adopting criterion (18), for $M_{\mathrm{S}} = 1$.

$$P_{\mathrm{SI},i} = \frac{P_{\mathrm{T}}}{K}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\int_{-\infty}^{+\infty}\boldsymbol{e}_i^H\boldsymbol{H}_{\mathrm{SI}}(f)\boldsymbol{F}[k]\boldsymbol{F}^H[k]\boldsymbol{H}_{\mathrm{SI}}^H(f)\boldsymbol{e}_i\varphi_k(f)\mathrm{d}f + \frac{P_{\mathrm{T}}}{K}\frac{1}{\eta_{\mathrm{T}}}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\int_{-\infty}^{+\infty}\boldsymbol{e}_i^H\boldsymbol{H}_{\mathrm{SI}}(f)\sum_{j=1}^{M_{\mathrm{T}}}(\boldsymbol{e}_j^H\boldsymbol{F}[k]\boldsymbol{F}^H[k]\boldsymbol{e}_j)\boldsymbol{e}_j\boldsymbol{e}_j^H\boldsymbol{H}_{\mathrm{SI}}^H(f)\boldsymbol{e}_i\mathrm{d}f \tag{55}$$

$$P_{\mathrm{SI},i} = \frac{P_{\mathrm{T}}}{K}\mathrm{tr}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\boldsymbol{F}^H[k]\underbrace{\left(\int_{-\infty}^{+\infty}\boldsymbol{H}_{\mathrm{SI}}^H(f)\boldsymbol{e}_i\boldsymbol{e}_i^H\boldsymbol{H}_{\mathrm{SI}}(f)\varphi_k(f)\mathrm{d}f\right)}_{\boldsymbol{C}_i[k]}\boldsymbol{F}[k] + \frac{P_{\mathrm{T}}}{K}\frac{1}{\eta_{\mathrm{T}}}\mathrm{tr}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\sum_{j=1}^{M_{\mathrm{T}}}\boldsymbol{F}^H[k]\underbrace{\int_{-\infty}^{+\infty}\left|\boldsymbol{e}_i^H\boldsymbol{H}_{\mathrm{SI}}(f)\boldsymbol{e}_j\right|^2\mathrm{d}f}_{g_{ij}}\boldsymbol{e}_j\boldsymbol{e}_j^H\boldsymbol{F}[k] \tag{56}$$

proposed, exhibiting different tradeoffs in terms of feasibility, complexity, achievable MI, and SIC performance.

APPENDIX A
PROOF OF LEMMA 1

Recalling (4) and (6), one has

$$\boldsymbol{S}_{\boldsymbol{s}_1}(f) = \frac{P_{\mathrm{T}}}{K}\sum_{k\in\mathcal{K}}\frac{1}{M_{\mathrm{S}}[k]}\boldsymbol{F}[k]\boldsymbol{F}^H[k]\varphi_k(f) + \boldsymbol{\Sigma}_{\mathrm{T}}, \quad (53)$$

where $\boldsymbol{\Sigma}_{\mathrm{T}}$ is given in (8). Substituting (53) into (11), and since

$$\mathrm{diag}\left(\boldsymbol{F}[k]\boldsymbol{F}^H[k]\right) = \sum_{j=1}^{M_{\mathrm{T}}}(\boldsymbol{e}_j^H\boldsymbol{F}[k]\boldsymbol{F}^H[k]\boldsymbol{e}_j)\boldsymbol{e}_j\boldsymbol{e}_j^H, \quad (54)$$

one obtains (55) at the bottom of the previous page. Using the cyclic property of the trace, (55) can be written as (56) at the bottom of the preceding page, where g_{ij} and $\boldsymbol{C}_i[k]$, as defined in Sec. II, are recognized. Rearranging (56), (14) follows.

APPENDIX B

Proof of Theorem 1

Since $\boldsymbol{X}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{T}}}\ \forall k \in \mathcal{K}$, tightness trivially holds if $M_{\mathrm{R}'} \geq M_{\mathrm{T}}$. Otherwise, consider the SVD $\boldsymbol{H}_1[k] = \widetilde{\boldsymbol{U}}[k]\boldsymbol{\Sigma}[k]\widetilde{\boldsymbol{V}}^H[k]$, with $\widetilde{\boldsymbol{V}}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times M_{\mathrm{R}'}}$. Let $\widetilde{\boldsymbol{W}}[k] \in \mathbb{C}^{M_{\mathrm{T}}\times(M_{\mathrm{T}}-M_{\mathrm{R}'})}$ be an orthonormal basis of the null space of $\boldsymbol{H}_1[k]$, so that $\boldsymbol{Q}[k] = [\,\widetilde{\boldsymbol{V}}[k]\ \widetilde{\boldsymbol{W}}[k]\,]$ is unitary, and let

$$\boldsymbol{Y}[k] \triangleq \boldsymbol{Q}^H[k]\boldsymbol{X}[k]\boldsymbol{Q}[k] = \begin{bmatrix} \boldsymbol{Y}_{11}[k] & \boldsymbol{Y}_{12}[k] \\ \boldsymbol{Y}_{12}^H[k] & \boldsymbol{Y}_{22}[k] \end{bmatrix}, \quad (57)$$

$$\overline{\boldsymbol{C}}_i[k] \triangleq \boldsymbol{Q}^H[k]\widetilde{\boldsymbol{C}}_i[k]\boldsymbol{Q}[k] = \begin{bmatrix} \overline{\boldsymbol{C}}_{i,11}[k] & \overline{\boldsymbol{C}}_{i,12}[k] \\ \overline{\boldsymbol{C}}_{i,12}^H[k] & \overline{\boldsymbol{C}}_{i,22}[k] \end{bmatrix}, \quad (58)$$

where $\boldsymbol{Y}_{11}[k] = \widetilde{\boldsymbol{V}}^H[k]\boldsymbol{X}[k]\widetilde{\boldsymbol{V}}[k] \in \mathbb{C}^{M_{\mathrm{R}'}\times M_{\mathrm{R}'}}$, $\boldsymbol{Y}_{22}[k] = \widetilde{\boldsymbol{W}}^H[k]\boldsymbol{X}[k]\widetilde{\boldsymbol{W}}[k] \in \mathbb{C}^{(M_{\mathrm{T}}-M_{\mathrm{R}'})\times(M_{\mathrm{T}}-M_{\mathrm{R}'})}$, and $\boldsymbol{Y}_{12}[k] = \widetilde{\boldsymbol{V}}^H[k]\boldsymbol{X}[k]\widetilde{\boldsymbol{W}}[k] \in \mathbb{C}^{M_{\mathrm{R}'}\times(M_{\mathrm{T}}-M_{\mathrm{R}'})}$, and analogously for the constituent blocks of the partition in (58). Regarding $\{\boldsymbol{Y}[k]\}$ as the optimization variables, constraints (24b)–(24c) read as

$$\text{(24b)} \ \Leftrightarrow\ \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{Y}_{11}[k] + \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{Y}_{22}[k] \leq K, \quad (59)$$

$$\text{(24c)} \ \Leftrightarrow\ \frac{1}{K}\sum_{k\in\mathcal{K}}\log\left|\boldsymbol{I}_{M_{\mathrm{R}'}} + \gamma\boldsymbol{\Sigma}^2[k]\boldsymbol{Y}_{11}[k]\right| \geq t. \quad (60)$$

Using the generalized Schur complement [69], (24d) reads as

$$\text{(24d)} \Leftrightarrow \begin{cases} \boldsymbol{Y}_{11}[k] \succeq \boldsymbol{0} \\ \boldsymbol{Y}_{22}[k] - \boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k] \succeq \boldsymbol{0} \\ \left(\boldsymbol{I}_{M_{\mathrm{R}'}} - \boldsymbol{Y}_{11}[k]\boldsymbol{Y}_{11}^\sharp[k]\right)\boldsymbol{Y}_{12}[k] = \boldsymbol{0} \end{cases} \quad \forall k \in \mathcal{K}. \quad (61)$$

Consider now the SDR problem derived from $\mathbb{P}_1$, *i.e.,* (24a)-(24d). For any given $\{\boldsymbol{Y}_{11}[k], \boldsymbol{Y}_{12}[k]\}_{k\in\mathcal{K}}$, let $\boldsymbol{\Delta}[k] = \boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k]$ and $T_{11} = \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{Y}_{11}[k]$. Then the optimal $\{\boldsymbol{Y}_{22}[k]\}_{k\in\mathcal{K}}$ are the solution to

$$\min_{\{\boldsymbol{Y}_{22}[k]\in\mathbb{H}^{M_{\mathrm{T}}-M_{\mathrm{R}'}}\}} \quad \mathrm{tr}\sum_{k\in\mathcal{K}}\sum_{i=1}^{M_{\mathrm{R}}}\overline{\boldsymbol{C}}_{i,22}[k]\boldsymbol{Y}_{22}[k] \quad (62a)$$

$$\text{s.t.} \quad \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{Y}_{22}[k] \leq K - T_{11}, \quad (62b)$$

$$\boldsymbol{Y}_{22}[k] - \boldsymbol{\Delta}[k] \succeq \boldsymbol{0},\ \forall k \in \mathcal{K}. \quad (62c)$$

Note now that, for any $\boldsymbol{A} \succeq \boldsymbol{0}$, $\boldsymbol{\Delta} \succeq \boldsymbol{0}$, both of size $n \times n$,

$$\boldsymbol{\Delta} = \arg\min_{\boldsymbol{Z}\in\mathbb{H}^n} \mathrm{tr}\boldsymbol{A}\boldsymbol{Z} \quad \text{s.t.} \quad \boldsymbol{Z} - \boldsymbol{\Delta} \succeq \boldsymbol{0}. \quad (63)$$

Using (63), it is seen that (62) is feasible iff $\mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{\Delta}[k] \leq K - T_{11}$, in which case the solution is $\boldsymbol{Y}_{22}[k] = \boldsymbol{\Delta}[k]$, $\forall k \in \mathcal{K}$. Now, the rank additivity property of the generalized Schur complement [69, Ch. 6] states that $\mathrm{rank}(\boldsymbol{Y}[k]) = \mathrm{rank}(\boldsymbol{Y}_{11}[k]) + \mathrm{rank}(\boldsymbol{Y}_{22}[k] - \boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k])$; hence, for $\boldsymbol{Y}_{22}[k] = \boldsymbol{\Delta}[k]$, one has $\mathrm{rank}(\boldsymbol{X}[k]) = \mathrm{rank}(\boldsymbol{Y}[k]) = \mathrm{rank}(\boldsymbol{Y}_{11}[k]) \leq M_{\mathrm{R}'}$ $\forall k \in \mathcal{K}$, so that the SDR approach is tight. The proof for the SDR problem derived from $\mathbb{P}_2$ is analogous, upon substituting the objective in (62a) by $\max_{1\leq i\leq M_{\mathrm{R}}}\left(c_i + \mathrm{tr}\sum_{k\in\mathcal{K}}\overline{\boldsymbol{C}}_{i,22}[k]\boldsymbol{Y}_{22}[k]\right)$, where $c_i = \mathrm{tr}\sum_{k\in\mathcal{K}}\left(\overline{\boldsymbol{C}}_{i,11}[k]\boldsymbol{Y}_{11}[k] + \overline{\boldsymbol{C}}_{i,12}[k]\boldsymbol{Y}_{12}^H[k] + \overline{\boldsymbol{C}}_{i,12}^H[k]\boldsymbol{Y}_{12}[k]\right)$.

Proof of Corollary 1

Consider the SDR problem derived from $\mathbb{P}_1$. For $p, q \in \{1, 2\}$, and with $\overline{\boldsymbol{C}}_{i,pq}[k]$ as in (58), let $\overline{\boldsymbol{C}}_{pq}[k] = \sum_{i=1}^{M_{\mathrm{R}}}\overline{\boldsymbol{C}}_{i,pq}[k]$. With the optimal value $\boldsymbol{Y}_{22}[k] = \boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k]$, consider now the optimization of $\{\boldsymbol{Y}_{12}[k]\}_{k\in\mathcal{K}}$, for given $\{\boldsymbol{Y}_{11}[k]\}_{k\in\mathcal{K}}$:

$$\min_{\{\boldsymbol{Y}_{12}[k]\}} \ \mathrm{tr}\sum_{k\in\mathcal{K}}\Big(\overline{\boldsymbol{C}}_{11}[k]\boldsymbol{Y}_{11}[k] + \overline{\boldsymbol{C}}_{22}[k]\boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k] + \overline{\boldsymbol{C}}_{12}[k]\boldsymbol{Y}_{12}^H[k] + \overline{\boldsymbol{C}}_{12}^H[k]\boldsymbol{Y}_{12}[k]\Big) \quad (64a)$$

$$\text{s.t.} \ \left(\boldsymbol{I}_{M_{\mathrm{R}'}} - \boldsymbol{Y}_{11}[k]\boldsymbol{Y}_{11}^\sharp[k]\right)\boldsymbol{Y}_{12}[k] = \boldsymbol{0}\ \ \forall k \in \mathcal{K}. \quad (64b)$$

From (64b), $\boldsymbol{Y}_{12}[k]$ must lie in the column space of $\boldsymbol{Y}_{11}[k]$, *i.e.,* $\boldsymbol{Y}_{12}[k] = \boldsymbol{Y}_{11}[k]\boldsymbol{Z}[k]$ for some $\boldsymbol{Z}[k] \in \mathbb{C}^{M_{\mathrm{R}'}\times(M_{\mathrm{T}}-M_{\mathrm{R}'})}$. With this, and noting that $\boldsymbol{Y}_{11}[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{11}[k] = \boldsymbol{Y}_{11}[k]$, (64) can be rewritten as

$$\min_{\{\boldsymbol{Z}[k]\}} \ \mathrm{tr}\sum_{k\in\mathcal{K}}\big(\overline{\boldsymbol{C}}_{22}[k]\boldsymbol{Z}^H[k]\boldsymbol{Y}_{11}[k]\boldsymbol{Z}[k] + \overline{\boldsymbol{C}}_{12}[k]\boldsymbol{Z}^H[k]\boldsymbol{Y}_{11}[k] + \overline{\boldsymbol{C}}_{12}^H[k]\boldsymbol{Y}_{11}[k]\boldsymbol{Z}[k]\Big), \quad (65)$$

which is a convex problem, whose first-order optimality condition reads as $\boldsymbol{Y}_{11}[k](\boldsymbol{Z}[k]\overline{\boldsymbol{C}}_{22}[k] + \overline{\boldsymbol{C}}_{12}[k]) = \boldsymbol{0}\ \forall k \in \mathcal{K}$. The solution is not unique if $\overline{\boldsymbol{C}}_{22}[k]$ is singular; a particular solution is given by $\boldsymbol{Z}[k] = -\overline{\boldsymbol{C}}_{12}[k]\overline{\boldsymbol{C}}_{22}^\sharp[k]\ \forall k \in \mathcal{K}$, so that

$$\boldsymbol{Y}_{12}[k] = -\boldsymbol{Y}_{11}[k]\overline{\boldsymbol{C}}_{12}[k]\overline{\boldsymbol{C}}_{22}^\sharp[k] \quad \forall k \in \mathcal{K}. \quad (66)$$

Using (66), the optimal $\boldsymbol{Y}_{22}[k] = \boldsymbol{Y}_{12}^H[k]\boldsymbol{Y}_{11}^\sharp[k]\boldsymbol{Y}_{12}[k]$ is

$$\boldsymbol{Y}_{22}[k] = \overline{\boldsymbol{C}}_{22}^\sharp[k]\overline{\boldsymbol{C}}_{12}^H[k]\boldsymbol{Y}_{11}[k]\overline{\boldsymbol{C}}_{12}[k]\overline{\boldsymbol{C}}_{22}^\sharp[k] \quad \forall k \in \mathcal{K}. \quad (67)$$

Thus, with (66)-(67), one is left to finding the $M_{\mathrm{R}'} \times M_{\mathrm{R}'}$ matrices $\{\boldsymbol{Y}_{11}[k]\}$. Upon defining, $\forall k \in \mathcal{K}$,

$$\boldsymbol{A}[k] = \overline{\boldsymbol{C}}_{11}[k] - \overline{\boldsymbol{C}}_{12}[k]\overline{\boldsymbol{C}}_{22}^\sharp[k]\overline{\boldsymbol{C}}_{12}^H[k], \quad (68)$$

$$\boldsymbol{B}[k] = \boldsymbol{I}_{M_{\mathrm{R}'}} + \overline{\boldsymbol{C}}_{12}[k](\overline{\boldsymbol{C}}_{22}^\sharp[k])^2\overline{\boldsymbol{C}}_{12}^H[k], \quad (69)$$

then one must solve the convex problem

$$\min_{\{\boldsymbol{Y}_{11}[k]\in\mathbb{H}^{M_{\mathrm{R}'}}\}} \mathrm{tr}\sum_{k\in\mathcal{K}}\boldsymbol{A}[k]\boldsymbol{Y}_{11}[k] \quad (70a)$$

$$\text{s.t. (60)},\ \text{tr}\sum_{k\in\mathcal{K}}\boldsymbol{B}[k]\boldsymbol{Y}_{11}[k]\leq K,\ \boldsymbol{Y}_{11}[k]\succeq\boldsymbol{0}\ \ \forall k\in\mathcal{K},\quad (70b)$$

after which the matrices $\{\boldsymbol{Y}[k]\}$ will be available. From them, $\boldsymbol{X}[k]=\boldsymbol{Q}[k]\boldsymbol{Y}[k]\boldsymbol{Q}^H[k]$ are recovered $\forall k\in\mathcal{K}$.

APPENDIX C
PROOF OF LEMMA 2

Let $\omega[k]=\gamma\omega_1[k]$. For the considered setting, $\mathbb{P}_1'$ becomes

$$\min_{\{\lambda[k]\}_{k\in\mathcal{K}}}\quad\sum_{k\in\mathcal{K}}\lambda[k]c[k]\quad(71a)$$

$$\text{s.t.}\quad\sum_{k\in\mathcal{K}}\lambda[k]\leq K,\quad(71b)$$

$$\frac{1}{K}\sum_{k\in\mathcal{K}}\log(1+\lambda[k]\omega[k])\geq t,\quad(71c)$$

$$\lambda[k]\geq 0,\ \forall k\in\mathcal{K}.\quad(71d)$$

The Lagrangian associated with (71) reads as

$$\mathcal{L}=\sum_{k\in\mathcal{K}}\lambda[k]c[k]+\mu\left(\sum_{k\in\mathcal{K}}\lambda[k]-K\right)+\chi\left(Kt-\sum_{k\in\mathcal{K}}\log(1+\lambda[k]\omega[k])\right)-\sum_{k\in\mathcal{K}}\eta_k\lambda[k],\quad(72)$$

with μ, χ, $\{\eta_k\}_{k\in\mathcal{K}}$ the Lagrange multipliers associated with constraints (71b), (71c), (71d), respectively. At the solution, (71c) must be met with equality, thus, $\chi>0$; otherwise, such solution could be scaled by $1-\epsilon$, with $\epsilon>0$ sufficiently small, still satisfying all constraints while reducing the objective.

Since $\{\eta_k\}_{k\in\mathcal{K}}$ are slack variables, the Karush-Kuhn-Tucker (KKT) conditions reveal that

$$\left.\frac{\partial\mathcal{L}}{\partial\lambda[k]}\right|_{\{\lambda^\star[k]\}}\begin{cases}=0 & \text{if } \lambda^\star[k]>0,\\ \geq 0 & \text{if } \lambda^\star[k]=0,\end{cases}\quad(73)$$

where $\{\lambda^\star[k]\}$ is the optimal solution to (71). Thus, from (73),

$$c[k]+\mu-\chi\frac{\omega[k]}{1+\omega[k]\lambda^\star[k]}\geq 0\quad\forall k\in\mathcal{K},\quad(74)$$

which must hold with equality to meet complementary slackness whenever $\lambda^\star[k]>0$. Thus,

$$\lambda^\star[k]=\begin{cases}\frac{\chi}{\mu+c[k]}-\frac{1}{\omega[k]} & \text{if } \mu+c[k]>\frac{\chi\omega[k]}{1+\omega[k]\lambda^\star[k]},\\ 0 & \text{otherwise,}\end{cases}\quad(75)$$

which amounts to (33).

REFERENCES

[1] B. Smida, A. Sabharwal, G. Fodor, G. C. Alexandropoulos, H. A. Suraweera, and C.-B. Chae, "Full-duplex wireless for 6G: Progress brings new opportunities and challenges," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2729–2750, 2023.

[2] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[3] K. Shahzad, X. Zhou, S. Yan, J. Hu, F. Shu, and J. Li, "Achieving covert wireless communications using a full-duplex receiver," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 8517–8530, 2018.

[4] D. Li, D. Zhang, and G. Zhang, "Degrees of freedom for half-duplex and full-duplex multi-user cognitive radios," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2812–2827, 2020.

[5] P. Dehghanzadeh, A. Madanayake, H. Zhao, S. B. Venkatakrishnan, and S. Mandal, "A multiport self-interference canceller for wideband SIMO/MIMO-STAR full-duplex arrays," *IEEE Trans. Microw. Theory Tech.*, pp. 1–15, 2023.

[6] Z. Zhou, Y. Li, J. Hu, Y. He, Z. Zhang, and P.-Y. Chen, "Monostatic copolarized simultaneous transmit and receive (STAR) antenna by integrated single-layer design," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 3, pp. 472–476, 2019.

[7] J. C. Dash and D. Sarkar, "A colinearly polarized full-duplex antenna with extremely high Tx–Rx isolation," *IEEE Antennas Wireless Propag. Lett.*, vol. 21, no. 12, pp. 2387–2391, 2022.

[8] M. N. A. Tarek, R. Hokayem, S. R. Govindarajulu, M. H. Novak, and E. A. Alwan, "A two-stage wideband RF cancellation of coupled transmit signal for bi-static simultaneous transmit and receive system," *IEEE J. Microw.*, vol. 2, no. 3, pp. 429–441, 2022.

[9] A. T. Le, L. C. Tran, X. Huang, and Y. J. Guo, "Beam-based analog self-interference cancellation in full-duplex MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 4, pp. 2460–2471, 2020.

[10] H. Luo, M. Holm, and T. Ratnarajah, "Wideband active analog self-interference cancellation for 5G and beyond full-duplex systems," in *2020 54th Asilomar Conf. Signals, Syst., Comput.*, 2020, pp. 868–872.

[11] J. W. Kwak, M. S. Sim, I.-W. Kang, J. Park, K.-K. Wong, and C.-B. Chae, "Analog self-interference cancellation with practical RF components for full-duplex radios," *IEEE Trans. Wireless Commun.*, vol. 22, no. 7, pp. 4552–4564, 2023.

[12] C. W. Morgenstern, Y. Rong, A. Herschfelt, A. C. Molnar, A. B. Apsel, D. G. Landon, and D. W. Bliss, "Analog-domain self-interference cancellation for practical multi-tap full-duplex system: Theory, modeling, and algorithm," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2796–2807, 2023.

[13] C. Zhang and X. Luo, "Adaptive digital self-interference cancellation for millimeter-wave full-duplex backhaul systems," *IEEE Access*, vol. 7, pp. 175 542–175 553, 2019.

[14] Y. He, H. Zhao, W. Guo, S. Shao, and Y. Tang, "A time-robust digital self-interference cancellation in full-duplex radios: Receiver design and performance analysis," *IEEE Access*, vol. 8, pp. 185 021–185 031, 2020.

[15] ——, "Frequency-domain successive cancellation of nonlinear self-interference with reduced complexity for full-duplex radios," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2678–2690, 2022.

[16] M. Mohammadi, Z. Mobini, D. Galappaththige, and C. Tellambura, "A comprehensive survey on full-duplex communication: Current solutions, future trends, and open issues," *IEEE Commun. Surveys Tut.*, vol. 25, no. 4, pp. 2190–2244, 2023.

[17] J. P. Doane, K. E. Kolodziej, and B. T. Perry, "Simultaneous transmit and receive with digital phased arrays," in *2016 IEEE Int. Symp. Phased Array Syst. Technol. (PAST)*, 2016, pp. 1–6.

[18] B. Smida, R. Wichman, K. E. Kolodziej, H. A. Suraweera, T. Riihonen, and A. Sabharwal, "In-band full-duplex: The physical layer," *Proc. IEEE*, vol. 112, no. 5, pp. 433–462, 2024.

[19] G. C. Alexandropoulos, M. A. Islam, and B. Smida, "Full-duplex massive multiple-input, multiple-output architectures: Recent advances, applications, and future directions," *IEEE Veh. Technol. Mag.*, vol. 17, no. 4, pp. 83–91, 2022.

[20] M. A. Islam, G. C. Alexandropoulos, and B. Smida, "Joint analog and digital transceiver design for wideband full duplex MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9729–9743, 2022.

[21] M. Erdem, O. Gurbuz, D. C. Altintas, and H. Ayar, "A switched full duplex MIMO architecture with digital linear and nonlinear cancellation," *Physical Commun.*, vol. 68, p. 102542, 2025.

[22] X. Chen, V. Savaux, M. Crussière, P. Savelli, and K.-C. Yao, "Energy-efficient spatial domain self-interference cancellation method by the addition of an RF chain in full-duplex MIMO systems," *IEEE Trans. Veh. Technol.*, 2025.

[23] E. Everett, C. Shepard, L. Zhong, and A. Sabharwal, "SoftNull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8077–8092, 2016.

[24] R. Sultan, K. G. Seddik, Z. Han, and B. Aazhang, "Joint transmitter-receiver optimization and self-interference suppression in full-duplex MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 6913–6929, 2021.

[25] S. E. Johnston and P. D. Fiore, "Full-duplex communication via adaptive nulling," in *2013 Asilomar Conf. Signals, Syst. Comput.*, 2013, pp. 1628–1631.

[26] B. Chun and H. Park, "A spatial-domain joint-nulling method of self-interference in full-duplex relays," *IEEE Commun. Lett.*, vol. 16, no. 4, pp. 436–438, 2012.

[27] M. A. Ahmed, C. C. Tsimenidis, and A. F. Al 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, 2016.

[28] N. M. Gowda and A. Sabharwal, "JointNull: Combining partial analog cancellation with transmit beamforming for large-antenna full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 2094–2108, 2018.

[29] I. T. Cummings, J. P. Doane, T. J. Schulz, and T. C. Havens, "Aperture-level simultaneous transmit and receive with digital phased arrays," *IEEE Trans. Signal Process.*, vol. 68, pp. 1243–1258, 2020.

[30] X. Huang, A. Tuyen Le, and Y. J. Guo, "Transmit beamforming for communication and self-interference cancellation in full duplex MIMO systems: A trade-off analysis," *IEEE Trans. Wireless Commun.*, vol. 20, no. 6, pp. 3760–3769, 2021.

[31] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, 2015.

[32] C. K. Sheemar and D. Slock, "Beamforming for bidirectional MIMO full duplex under the joint sum power and per antenna power constraints," in *IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2021, pp. 4800–4804.

[33] C. K. Sheemar, C. K. Thomas, and D. Slock, "Practical hybrid beamforming for millimeter wave massive MIMO full duplex with limited dynamic range," *IEEE Open J. Commun. Soc.*, vol. 3, pp. 127–143, 2022.

[34] Z. Jiang, L. Xia, F. Xu, C. Liu, and H. Long, "Joint transmitter and receiver precoding for frequency-selective sub-band full-duplex systems," in *GLOBECOM 2025-2025 IEEE Global Commun. Conf.* IEEE, 2025, pp. 5067–5072.

[35] C. Shi, W. Pan, Y. Shen, and S. Shao, "Robust transmit beamforming for self-interference cancellation in STAR phased array systems," *IEEE Signal Process. Lett.*, vol. 29, pp. 2622–2626, 2022.

[36] A. Hakimi, S. Zargari, C. Tellambura, and S. Herath, "Sum rate maximization of MIMO monostatic backscatter networks by suppressing residual self-interference," *IEEE Trans. Commun.*, vol. 71, no. 1, pp. 512–526, 2022.

[37] C. Shi, W. Pan, Y. Shen, and S. Shao, "Transmit beamforming self-interference cancellation design and experiment for STAR phased array," in *2024 IEEE Globecom Workshops*. IEEE, 2024, pp. 1–6.

[38] X. Chen, V. Savaux, M. Crussière, P. Savelli, and K. C. Yao, "Self-interference cancellation for MIMO full-duplex downlink systems: A constrained MMSE-based precoding approach," *IEEE Access*, 2025.

[39] A. Liu, W. Sheng, and T. Riihonen, "Per-antenna self-interference cancellation beamforming design for digital phased array," *IEEE Signal Process. Lett.*, vol. 29, pp. 2442–2446, 2022.

[40] B. P. Day, A. R. Margetts, D. W. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3702–3713, 2012.

[41] I. P. Roberts, J. G. Andrews, and S. Vishwanath, "Hybrid beamforming for millimeter wave full-duplex under limited receive dynamic range," *IEEE Trans. Wireless Commun.*, vol. 20, no. 12, pp. 7758–7772, 2021.

[42] A. C. Cirik, L. Zhou, and T. Ratnarajah, "Linear transceiver design with per-antenna power constraints in full-duplex multi-user MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 4, pp. 412–415, 2016.

[43] O. Taghizadeh, V. Radhakrishnan, A. C. Cirik, R. Mathar, and L. Lampe, "Hardware impairments aware transceiver design for bidirectional full-duplex MIMO OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7450–7464, 2018.

[44] H. Shen, C. Liu, W. Xu, and C. Zhao, "Optimized full-duplex MIMO DF relaying with limited dynamic range," *IEEE Access*, vol. 5, pp. 20 726–20 735, 2017.

[45] V. Radhakrishnan, O. Taghizadeh, and R. Mathar, "Hardware impairments-aware transceiver design for multi-carrier full-duplex MIMO relaying," *IEEE Trans. Veh. Technol.*, vol. 70, no. 2, pp. 1109–1121, 2021.

[46] J. Jin, Y. R. Zheng, W. Chen, and C. Xiao, "Generalized quadratic matrix programming: A unified framework for linear precoding with arbitrary input distributions," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4887–4901, 2017.

[47] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, 2010.

[48] R. Sultan, K. G. Seddik, Z. Han, and B. Aazhang, "Joint transmitter-receiver optimization and self-interference suppression in full-duplex MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 6913–6929, 2021.

[49] M. A. Islam, G. C. Alexandropoulos, and B. Smida, "Joint analog and digital transceiver design for wideband full duplex MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 9729–9743, 2022.

[50] C. B. Barneto, T. Riihonen, S. D. Liyanaarachchi, M. Heino, N. González-Prelcic, and M. Valkama, "Beamformer design and optimization for joint communication and full-duplex sensing at mm-Waves," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8298–8312, 2022.

[51] A. Masmoudi and T. Le-Ngoc, "A maximum-likelihood channel estimator for self-interference cancelation in full-duplex systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 7, pp. 5122–5132, 2016.

[52] ——, "Channel estimation and self-interference cancelation in full-duplex communication systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 1, pp. 321–334, 2017.

[53] A. M. M. Chandran, L. Wang, and M. Zawodniok, "Channel estimators for full-duplex communication using orthogonal pilot sequences," *IEEE Access*, vol. 8, pp. 117 706–117 713, 2020.

[54] G. Santella and F. Mazzenga, "A hybrid analytical-simulation procedure for performance evaluation in M-QAM-OFDM schemes in presence of nonlinear distortions," *IEEE Trans. Veh. Technol.*, vol. 47, no. 1, pp. 142–151, 1998.

[55] W. Namgoong, "Modeling and analysis of nonlinearities and mismatches in AC-coupled direct-conversion receiver," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 163–173, 2005.

[56] 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, pp. 867–876, 2008.

[57] 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, 2012.

[58] R. López-Valcarce, "General form of the power spectral density of multicarrier signals," *IEEE Commun. Lett.*, vol. 26, no. 8, pp. 1755–1759, 2022.

[59] G. Pataki, *The Geometry of Semidefinite Programming*. Boston, MA: Springer US, 2000, pp. 29–65.

[60] A. Beck, "Quadratic matrix programming," *SIAM J. Optim.*, vol. 17, no. 4, pp. 1224–1238, 2007.

[61] A. S. Nemirovski and M. J. Todd, "Interior-point methods for optimization," *Acta Numerica*, vol. 17, pp. 191–234, 2008.

[62] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM review*, vol. 38, no. 1, pp. 49–95, 1996.

[63] 3GPP, "5G; NR; Physical layer procedures for data (Release 18)," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.214, May 2024, version 18.2.0.

[64] C. Xing, Y. Jing, S. Wang, S. Ma, and H. V. Poor, "New viewpoint and algorithms for water-filling solutions in wireless communications," *IEEE Trans. Signal Process.*, vol. 68, pp. 1618–1634, 2020.

[65] S. Kang, M. Mezzavilla, S. Rangan, A. Madanayake, S. B. Venkatakrishnan, G. Hellbourg, M. Ghosh, H. Rahmani, and A. Dhananjay, "Cellular wireless networks in the upper mid-band," *IEEE Open J. Commun. Soc.*, vol. 5, pp. 2058–2075, 2024.

[66] R. López-Valcarce and M. Martínez-Cotelo, "Full-duplex mmWave MIMO with finite-resolution phase shifters," *IEEE Trans. Wireless Commun.*, vol. 21, no. 11, pp. 8979–8994, 2022.

[67] Z. Xiao, P. Xia, and X.-G. Xia, "Full-duplex millimeter-wave communication," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 136–143, 2017.

[68] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," https://cvxr.com/cvx, Mar. 2014.

[69] F. Zhang, *The Schur complement and its applications*. Springer Science & Business Media, 2006, vol. 4.