Robust Transceiver Design in Full-Duplex MIMO Cognitive Radios

Robust Transceiver Design in Full-Duplex MIMO Cognitive Radios

Ali Cagatay Cirik, *Member, IEEE*, Sudip Biswas, *Member, IEEE*,
Omid Taghizadeh, *Student Member, IEEE* and Tharmalingam Ratnarajah, *Senior Member, IEEE*

***Abstract*—We consider a full duplex (FD) multiple-input multiple-output (MIMO) underlay cognitive radio (CR) cellular network, in which an FD secondary base-station (BS) serves multiple half-duplex (HD) uplink (UL) and downlink (DL) secondary users (SUs) at the same time and frequency. We assume that the channel state information (CSI) available at the transmitters is imperfect, and the errors of the CSI are assumed to be norm bounded. Under the impact of channel uncertainty, we address the sum mean-squared-errors (MSE) minimization problem subject to individual power constraints at the UL SUs, a total power-constraint at the secondary BS, and the interference constraints on the primary users (PUs) by the secondary network. By transforming the problem into an equivalent semidefinite programming (SDP), we propose an iterative alternating algorithm to compute the transceiver matrices jointly. Moreover, to reduce the high computational complexity of the SDP method, we develop a cutting-set method, which solves the problem by alternating between an optimization step (transceiver design) and a pessimization step (worst-case channel analysis). Numerical results are presented to show the effectiveness and robustness of the proposed algorithms.**

***Keywords*—Cognitive radio, full-duplex, MIMO, MSE, multi-user, self-interference, robust transceiver designs.**

I. Introduction

With the dramatic increase in the demand for wireless communication services, the efficient utilization of the available spectrum resources becomes crucial, and motivates the realization of new solutions. In particular, a full-duplex (FD) transceiver is able to transmit and receive at the same time and the same frequency, and hence has the potential to enhance the spectral efficiency. As a result it is widely considered as a candidate technology for future communication systems [2]. Nevertheless, FD systems suffer from the inherent interference from their own transmit signal, i.e., self-interference (SI). In order to avoid the SI, the currently deployed wireless communication systems operate in half-duplex (HD) mode by allocating orthogonal channels for transmission and reception. This is usually realized by separating the channels via the well known time division duplexing (TDD), or frequency division duplexing (FDD) schemes.

Recently, it has been experimentally demonstrated that the overwhelming SI in an FD transceiver can be sufficiently mitigated by simultaneously implementing cancellation schemes including antenna design, and analog and digital domain SI cancellation techniques [3]-[5]. However, consideration of a residual SI is still necessary, due to the inherent imperfection of the transmit and receive chains [6]-[13]. Moreover, the interference from the uplink (UL) to downlink (DL) users, i.e., the co-channel interference (CCI), is another challenge for the gainful use of FD technology. In this respect, the application of beamforming techniques are known to be effective, where the impact of the CCI, as well as the SI are jointly taken into consideration [14], [15].

Furthermore, cognitive radio (CR) is another promising technology that can enhance spectrum efficiency [16], [17]. In an underlay CR system, unlicensed secondary users (SUs) can access the spectrum owned by the licensed primary users (PUs) as long as the interference level from SUs to PUs is maintained to meet the Quality-of-Service (QoS) requirements for the PUs. In practice, it is quite difficult to obtain the estimates of the channels in CR systems due to the time varying and random nature of wireless channel, and the lack of full SU-PU communication. In this regard, there are two approaches which are commonly used to model the channel state information (CSI) error. Firstly, the stochastic approach, where the channel is usually modeled as a complex random matrix with normally distributed elements, and the transmitter knows the mean and/or the covariance, i.e., the slowly-varying channel statistics that can be well estimated [18]. Secondly, the deterministic (or worst-case) approach, where the instantaneous CSI error is assumed to lie in a known set, e.g., a norm-2 ball, where a larger set represents a higher CSI uncertainty [19]-[23].

The authors in [12] studied the minimization of sum mean-squared-error (MSE) in a multiple-input multiple-output (MIMO) FD underlay CR system. By assuming perfect knowledge of the CSI, a second-order cone programming (SOCP)-based algorithm was proposed. However, the sum-MSE optimization problem under norm-bounded imperfect CSI cannot be cast as an SOCP problem, and thus the algorithms proposed in [12] are not applicable under the assumption of imperfect CSI knowledge, since it is not suitable to solve the min-max problems arising from worst-case (norm-bounded) CSI considered in this paper. Hence, it is of paramount importance to consider the imperfect channel estimates, and develop robust beamforming schemes in FD underlay CR systems.

Motivated by the above, in this paper, we study an underlay CR system, where a secondary FD base-station (BS) serves multiple HD UL and DL SUs at the same time and frequency, and the secondary network utilizes the spectrum simultaneously with multiple PUs. In addition to SI, CCI is also

A. C. Cirik was with the School of Engineering, Institute for Digital Communications, University of Edinburgh, Edinburgh EH9 3JL, U.K. He is now with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada (email: cirik@ece.ubc.ca).

S. Biswas and T. Ratnarajah are with the School of Engineering, Institute for Digital Communications, University of Edinburgh, Edinburgh EH9 3JL, U.K. (email: {sudip.biswas, t.ratnarajah}@ed.ac.uk).

O. Taghizadeh is with the Institute for Theoretical Information Technology, RWTH Aachen University, Aachen, 52074, Germany (email: taghizadeh@ti.rwth-aachen.de).

This work was supported by the Seventh Framework Programme for Research of the European Commission under grant number ADEL - 619647 and the U.K. Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/L025299/1.

Part of this work was presented in IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, May 2016 [1].

Correspondent author: Sudip Biswas.

taken into account to design the optimal robust beamformers under a norm-bounded-error model. We study the sum-MSE minimization problem under the power constraints at the UL SUs and secondary BS, and SUs' interference constraints to the PUs. Since this problem is semi-infinite (optimization problems with infinitely many constraints) [24, Ch. 3], we propose two methods to design the transceiver matrices at the secondary network jointly. In the first method, by representing the semi-infinite constraints in tractable forms (linear matrix inequality (LMI)), the problem can be transformed to a semidefinite programming (SDP), and therefore efficiently solved by optimizing the transmit and receiving beamforming matrices iteratively. Due to the prohibitive complexity of the SDP-based algorithms, in the second algorithm, we propose a low complexity cutting-set method [25], [26], which solves the problem by alternating between an optimization step and a pessimization step. In the first (optimization) step, the transmit and receive beamforming matrices are designed under a given finite subset of the uncertainty region, and in the second (pessimization) step, the subset is updated through the computation of the worst-case channels in the uncertainty regions.

A. Related Works

The imperfect CSI in FD cellular systems have been considered in the beamformer design in [27], [28], [29], where the channel estimation errors are modeled using a statistical approach. Norm-bounded CSI model in FD systems was first adopted in [30], [31]. However, only single-users equipped with single antennas was considered. Therefore, they cannot be applied to FD multi-user MIMO wireless communication systems directly.

In the context of robust CR networks design, a stochastic model was considered in [18] and the worst-case model was considered in [19], [21], [22], [23]. For example, in [19], a robust cognitive beamforming scheme was proposed where the secondary transmitters were assumed to have the the perfect CSI information between secondary transmitters and receivers, whereas the channel between the secondary transmitters and primary users were assumed to be known imperfectly. Moreover, in [19], the transmitters and receivers were assumed to be equipped with multiple and single antennas, respectively leading to a convex optimization problem. In [21], the non-cooperative game was considered to optimize the CR network where the resource allocation problem for each game player was convex. Compared with the optimization problems in [19], [21], which are convex, the robust beamforming problem we have is more challenging because it is non-convex. Both our proposed scheme and the scheme proposed in [22] employ the S-procedure [32] to transform the constraints into LMIs. The difference between our proposed scheme and the scheme proposed in [22] is that in our proposed scheme, the optimization problem is transformed to a single SDP while in [22], the optimization problem is solved using bisection search where in each iteration an SDP is solved. Moreover, single-antenna receivers are assumed in [22]. In [23], each semi-infinite constraint included only one uncertainty variable, and they mostly resort to the complex-valued version of the sign-definiteness lemma published in [32] to resolve the semi-infiniteness of the constraints. We apply the sign-definiteness lemma to the case of complex-valued quantities with multiple uncertainties in each design constraint. Moreover, in [23], identical channel estimation errors for different secondary and primary links were assumed. Compared to [23], the proposed scheme in our paper accounts for different estimation inaccuracies in the secondary and primary links.

As discussed in [12], MSE-based beamforming problems have been considered extensively for many communication systems in literature due to their good performance, significantly reduced complexity, and relationship with bit error rate (BER) and signal-to-interference-plus-noise ratio (SINR). Here, we reveal useful insights into FD CR systems via a robust MSE-based optimization. Indeed, the studied system in our paper shares many similarities to the traditional multi-cell multi-user networks, both in terms of the system concept, service requirements, as well as the design guidelines. Nevertheless the application of the FD technology in a CR network introduces fundamental new challenges to the traditional multi-user MIMO systems, which is the main focus of the present paper. The main distinctions are summarized as follows:

- The coexistence requirements of a secondary cellular system, with a network of PUs needs to be revised for an FD CR system. This is because, in an FD cellular network, the BS, as well as the UL users share the same channel resource for transmission, which results in the imposition of a higher interference intensity on the primary network. This issue becomes more critical considering the fact that the acquisition of an accurate CSI regarding the interference paths from the secondary UL users to the primary nodes is relatively unrealistic, and calls for the consideration of joint robust transmission strategies from the UL secondary users as well as the secondary BS.
- In an FD multi-user MIMO system, the interference paths between UL and DL SUs should be additionally taken into account. This impacts both the system performance, as well as the design strategy.
- In an FD multi-user MIMO system, the SI at the secondary BS is a critical challenge, and strongly relates the performance/design of the UL reception to the DL transmission. In this respect, the consideration of an accurate transceiver model, including the impacts of transmission and reception distortions are critical, as it is well-established in the context of FD system design and analysis.

Note that the aforementioned considerations regarding the design of a robust FD multi-user MIMO CR network, result in a relatively complicated problem structure. Moreover, due to the transmit and receive distortions at the FD nodes in addition to imperfect CSI, MSE is a complicated function, which makes the transformation of the constraints in the optimization problems complicated. This, in turn, calls for a rigorous optimization and analysis, together with a dedicated computational complexity study. Simulation results demonstrate that the proposed FD system can achieve a significant improvement of throughput over HD system.

B. Notation

Matrices and vectors are denoted as bold capital and lowercase letters, respectively. $(\cdot)^T$ is the transpose; $(\cdot)^*$ is the conjugate, and $(\cdot)^H$ is the conjugate transpose. $\mathbb{E}\{\cdot\}$ and $\mathrm{tr}\{\cdot\}$ denote the statistical expectation and trace, respectively. $\mathbf{I}_N$ and $\mathbf{0}_{N\times M}$ denote the N by N identity and the N by M zero matrices, respectively. $\mathrm{diag}\left(\mathbf{A}\right)$ is the diagonal matrix with the same diagonal elements as $\mathbf{A}$. $\mathcal{CN}\left(\mu,\sigma^2\right)$ denotes a complex Gaussian distribution with mean μ and variance σ^2. $vec(\cdot)$ stacks the elements of a matrix to one long column vector.

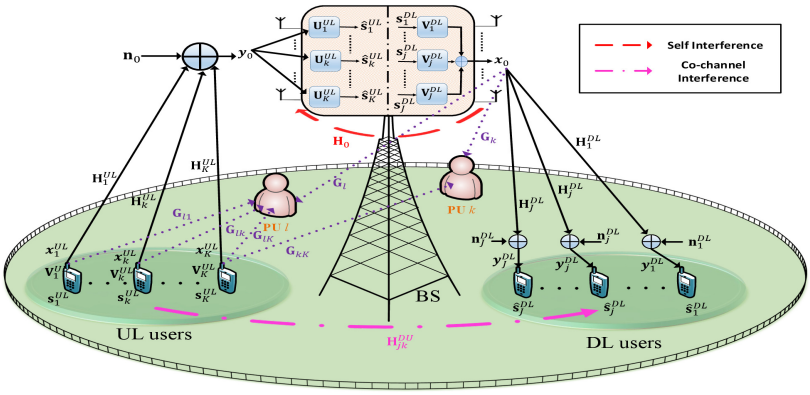


Fig. 1: An illustration of a full-duplex multi-user MIMO CR system model. The PUs are being interfered by both the SU BS and UL SUs.

The operators $\otimes$ and $\perp$ denote Kronecker product and the statistical independence, respectively. $\|\mathbf{X}\|_F$ and $\|\mathbf{x}\|_2$ denote the Frobenius norm of a matrix $\mathbf{X}$ and the Euclidean norm of a vector $\mathbf{x}$, respectively. $\lfloor \mathbf{A}_i \rfloor_{i=1,\ldots,K}$ denotes a tall matrix (or vector) obtained by stacking the matrices $\mathbf{A}_i$, $i = 1, \ldots, K$. $\Re\{\mathbf{A}\}$ represents the real part of $\mathbf{A}$. $\mathbf{A} \succeq \mathbf{0}$ shows that $\mathbf{A}$ is a positive semidefinite matrix. Finally, $\mathcal{R}^{m\times n}$ and $\mathbb{C}^{m\times n}$ denote $m \times n$ real and complex matrices, respectively.

II. System Model

We study an underlay FD CR cellular system, in which a secondary FD BS equipped with M_0 transmit and N_0 receive antennas serves K HD mode UL SUs, each equipped with M_k transmit antennas and J HD mode DL SUs, each equipped with N_j receive antennas simultaneously. The secondary network coexists with the primary network, consisting of L PUs as seen in Fig. 1.

In Fig. 1, $\mathbf{H}_k^{UL} \in \mathbb{C}^{N_0\times M_k}$ and $\mathbf{H}_j^{DL} \in \mathbb{C}^{N_j\times M_0}$ denote the k-th UL and the j-th DL channel, respectively. Moreover, $\mathbf{H}_0 \in \mathbb{C}^{N_0\times M_0}$ and $\mathbf{H}_{jk}^{DU} \in \mathbb{C}^{N_j\times M_k}$ denote the SI channel at the FD BS and the CCI channel between the k-th UL and j-th DL users, respectively.

We adopt the transmitter/receiver distortion model in [7], which accounts for the non-ideal hardware components at the transmit and receive chains, e.g., power amplifiers, oscillators, analog-to-digital converters (ADCs), and digital-to-analog converters (DACs). The aforementioned model relies on the experimental measurements in [33] and [34] showing that a collective effects of hardware imperfections can be approximated as an additive Gaussian noise, and this model has been commonly used in related literature, see [8]-[13]. In this respect, we consider an additive white Gaussian term as "transmitter noise" ("receiver distortion") at each transmit (receive) antenna, whose variance is κ (β) times the power of the undistorted signal at the corresponding chain.

The source symbols at the k-th UL and j-th DL users are denoted as $\mathbf{s}_k^{UL} \in \mathbb{C}^{d_k^{UL}\times 1}$ and $\mathbf{s}_j^{DL} \in \mathbb{C}^{d_j^{DL}\times 1}$, respectively. It is assumed that the symbols are independent and identically distributed (i.i.d.) with unit power, i.e., $\mathbb{E}\left[\mathbf{s}_k^{UL}\left(\mathbf{s}_k^{UL}\right)^H\right] = \mathbf{I}_{d_k^{UL}}$ and $\mathbb{E}\left[\mathbf{s}_j^{DL}\left(\mathbf{s}_j^{DL}\right)^H\right] = \mathbf{I}_{d_j^{DL}}$. The transmitted signal of the k-th UL user and that of the BS can be written as

$$\mathbf{x}_k^{UL} = \mathbf{V}_k^{UL}\mathbf{s}_k^{UL}, \qquad \mathbf{x}_0 = \sum_{j=1}^{J}\mathbf{V}_j^{DL}\mathbf{s}_j^{DL}, \tag{1}$$

respectively, where $\mathbf{V}_k^{UL} \in \mathbb{C}^{M_k\times d_k^{UL}}$ and $\mathbf{V}_j^{DL} \in \mathbb{C}^{M_0\times d_j^{DL}}$ represent the precoders for the data streams of the k-th UL and j-th DL users, respectively. The signal received by the BS and that received by the j-th DL user can be written as

$$\mathbf{y}_0 = \sum_{k=1}^{K}\mathbf{H}_k^{UL}\left(\mathbf{x}_k^{UL} + \mathbf{c}_k^{UL}\right) + \mathbf{H}_0\left(\mathbf{x}_0 + \mathbf{c}_0\right) + \mathbf{e}_0 + \mathbf{n}_0, \tag{2}$$

$$\mathbf{y}_j^{DL} = \mathbf{H}_j^{DL}\left(\mathbf{x}_0 + \mathbf{c}_0\right) + \sum_{k=1}^{K}\mathbf{H}_{jk}^{DU}\left(\mathbf{x}_k^{UL} + \mathbf{c}_k^{UL}\right) + \mathbf{e}_j^{DL} + \mathbf{n}_j^{DL}, \tag{3}$$

respectively, where $\mathbf{n}_0 \in \mathbb{C}^{N_0}$ and $\mathbf{n}_j^{DL} \in \mathbb{C}^{N_j}$ denote the additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $\mathbf{R}_0 = \sigma_0^2\mathbf{I}_{N_0}$ and $\mathbf{R}_j^{DL} = \sigma_j^2\mathbf{I}_{N_j}$ at the BS and the j-th DL user, respectively.[1]

Furthermore, $\mathbf{c}_k^{UL}$ ($\mathbf{c}_0$) in (2)-(3), is the distortion at the transmitter at the k-th UL user (BS), which closely approximates the effects of phase noise, non-linearities in the DAC and additive power-amplifier noise. The covariance matrix of $\mathbf{c}_k^{UL}$ is given by κ ($\kappa \ll 1$) times the energy of the intended signal at each transmit antenna [7]. In particular $\mathbf{c}_k^{UL}$ can be modeled as

$$\mathbf{c}_k^{UL} \sim \mathcal{CN}\left(\mathbf{0}, \kappa\,\mathrm{diag}\left(\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H\right)\right), \ \mathbf{c}_k^{UL} \perp \mathbf{x}_k^{UL}. \tag{4}$$

Finally, in (3)((2)), $\mathbf{e}_j^{DL}$ ($\mathbf{e}_0$) is the receiver distortion at the j-th DL user (BS), which closely approximates the combined effects of non-linearities in the ADC, additive gain-control noise and phase noise. The covariance matrix of $\mathbf{e}_j^{DL}$ is given by β ($\beta \ll 1$) times the energy of the undistorted received signal at each receive antenna [7] and can be modeled as

$$\mathbf{e}_j^{DL} \sim \mathcal{CN}\left(\mathbf{0}, \beta\mathrm{diag}\left(\mathbf{\Phi}_j^{DL}\right)\right), \ \mathbf{e}_j^{DL} \perp \mathbf{u}_j^{DL}, \tag{5}$$

where $\mathbf{\Phi}_j^{DL} = \mathrm{Cov}\{\mathbf{u}_j^{DL}\}$ and $\mathbf{u}_j^{DL}$ is the undistorted received vector at the j-th DL user, i.e., $\mathbf{u}_j^{DL} = \mathbf{y}_j^{DL} - \mathbf{e}_j^{DL}$. Similarly, the discussion on the transmitter/receiver distortion model holds for $\mathbf{c}_0$ and $\mathbf{e}_0$, as well.

To obtain the source symbols of the k-th UL and the j-th DL users, the received signals are multiplied by linear

[1] Since the secondary BS and DL SUs are unable to distinguish the interference from the primary network from the thermal noise, the noise terms in (2) and (3) include both the thermal noise and the interference from the PUs. This assumption is also adopted in [19], [21], [35]-[37] and the noise is modeled as zero mean with unit variance in [19], [38].

matrices, employing single-stream decoding[2]. In this regard, linear decoders are denoted as $\mathbf{U}_k^{UL} \in \mathbb{C}^{N_0 \times d_k^{UL}}$, and $\mathbf{U}_j^{DL} \in \mathbb{C}^{N_j \times d_j^{DL}}$ at the BS and at the j-th DL user, respectively. Consequently, the estimates of data streams are given as

$$\hat{\mathbf{s}}_k^{UL} = \left(\mathbf{U}_k^{UL}\right)^H \mathbf{y}_0, \qquad \hat{\mathbf{s}}_j^{DL} = \left(\mathbf{U}_j^{DL}\right)^H \mathbf{y}_j^{DL}. \tag{6}$$

Using these estimates, the MSEs of the k-th UL and j-th DL users can be written as in (7) and (8) given at the bottom of the current page, respectively. In (7) and (8), $\mathbf{\Sigma}_k^{UL}$ and $\mathbf{\Sigma}_j^{DL}$ are the approximated aggregate interference-plus-noise terms at the k-th UL and j-th DL user, respectively, and are expressed as in (9)-(10) given at the bottom of the following page[3] [4]. Since the secondary and primary networks coexist under the same spectrum, secondary network infers interference on the primary network. The interference power from the UL SUs and BS projected to the l-th PU equipped with T_l receive antennas is written as

$$I_l^{PU} = \sum_{k=1}^{K} \mathrm{tr}\Big\{\mathbf{G}_{lk}\Big(\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H + \kappa \mathrm{diag}\Big(\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H\Big)\Big)\mathbf{G}_{lk}^H\Big\} + \sum_{j=1}^{J} \mathrm{tr}\Big\{\mathbf{G}_l\Big(\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H + \kappa \mathrm{diag}\Big(\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H\Big)\Big)\mathbf{G}_l^H\Big\}, \tag{11}$$

where $\mathbf{G}_{lk} \in \mathbb{C}^{T_l \times M_k}$ $\left(\mathbf{G}_l \in \mathbb{C}^{T_l \times M_0}\right)$ is the channel between the l-th PU and k-th UL user (l-th PU and the BS).

A. Joint Beamforming Design

In this paper, we tackle the sum-MSE minimization problem, which is formulated as

$$\min_{\mathbf{V},\mathbf{U}} \quad \sum_{k=1}^{K} \mathrm{tr}\left\{\mathbf{MSE}_k^{UL}\right\} + \sum_{j=1}^{J} \mathrm{tr}\left\{\mathbf{MSE}_j^{DL}\right\} \tag{12a}$$

$$\text{s.t.} \quad \mathrm{tr}\left\{\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H\right\} \le P_k,\ k = 1, \ldots, K, \tag{12b}$$

$$\sum_{j=1}^{J} \mathrm{tr}\left\{\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H\right\} \le P_0, \tag{12c}$$

$$I_l^{PU} \le \lambda_l,\ l = 1, \ldots, L, \tag{12d}$$

where P_k in (12b) is the transmit power constraint at the k-th UL user, P_0 in (12c) is the total power constraint at the BS, and λ_l in (12d) is the upper limit of the interference allowed to be imposed on the l-th PU. Here $\mathbf{V} = \left\{\mathbf{V}_k^{UL},\ k = 1, \ldots, K,\ \mathbf{V}_j^{DL},\ j = 1, \ldots, J\right\}$ and $\mathbf{U} = \left\{\mathbf{U}_k^{UL},\ k = 1, \ldots, K,\ \mathbf{U}_j^{DL},\ j = 1, \ldots, J\right\}$ are the set of all transmit and receive beamforming matrices, respectively.

1) Simplification of Notations: To simplify the notations, we combine UL and DL channels, similar to [15]. Let us denote $\mathcal{S}^{UL}$ and $\mathcal{S}^{DL}$ to represent the set of K UL and J DL channels, respectively. Denoting $\mathbf{H}_{ij}$, $\mathbf{n}_i$, $\mathbf{G}_{lj}$ and receive (transmit) antenna numbers $\tilde{N}_i(\tilde{M}_i)$ as

$$\mathbf{H}_{ij} = \begin{cases} \mathbf{H}_j^{UL},\ i \in \mathcal{S}^{UL},\ j \in \mathcal{S}^{UL}, \\ \mathbf{H}_0,\ \ i \in \mathcal{S}^{UL},\ j \in \mathcal{S}^{DL}, \\ \mathbf{H}_{ij}^{DU},\ i \in \mathcal{S}^{DL},\ j \in \mathcal{S}^{UL}, \\ \mathbf{H}_i^{DL},\ i \in \mathcal{S}^{DL},\ j \in \mathcal{S}^{DL}, \end{cases} \mathbf{n}_i = \begin{cases} \mathbf{n}_0, & i \in \mathcal{S}^{UL}, \\ \mathbf{n}_i^{DL}, & i \in \mathcal{S}^{DL}, \end{cases}$$

$$\mathbf{G}_{lj} = \begin{cases} \mathbf{G}_{lj}, & j \in \mathcal{S}^{UL}, \\ \mathbf{G}_l, & j \in \mathcal{S}^{DL}, \end{cases} \tilde{N}_i\left(\tilde{M}_i\right) = \begin{cases} N_0\left(M_i\right), & i \in \mathcal{S}^{UL}, \\ N_i\left(M_0\right), & i \in \mathcal{S}^{DL}, \end{cases}$$

and referring to $\mathbf{V}_i^X$, $\mathbf{U}_i^X$, d_i^X and $\mathbf{\Sigma}_i^X$, $X \in \{UL, DL\}$ as $\mathbf{V}_i$, $\mathbf{U}_i$, d_i and $\mathbf{\Sigma}_i$, respectively, the MSE of the i-th link,

$$\mathbf{MSE}_k^{UL} = \left(\left(\mathbf{U}_k^{UL}\right)^H \mathbf{H}_k^{UL}\mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}}\right)\left(\left(\mathbf{U}_k^{UL}\right)^H \mathbf{H}_k^{UL}\mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}}\right)^H + \left(\mathbf{U}_k^{UL}\right)^H \mathbf{\Sigma}_k^{UL}\mathbf{U}_k^{UL}, \tag{7}$$

$$\mathbf{MSE}_j^{DL} = \left(\left(\mathbf{U}_j^{DL}\right)^H \mathbf{H}_j^{DL}\mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}}\right)\left(\left(\mathbf{U}_j^{DL}\right)^H \mathbf{H}_j^{DL}\mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}}\right)^H + \left(\mathbf{U}_j^{DL}\right)^H \mathbf{\Sigma}_j^{DL}\mathbf{U}_j^{DL}. \tag{8}$$

$$\begin{aligned} \mathbf{\Sigma}_k^{UL} \approx & \sum_{j \ne k}^{K} \mathbf{H}_j^{UL}\mathbf{V}_j^{UL}\left(\mathbf{V}_j^{UL}\right)^H\left(\mathbf{H}_j^{UL}\right)^H + \kappa \sum_{j=1}^{K} \mathbf{H}_j^{UL}\mathrm{diag}\left(\mathbf{V}_j^{UL}\left(\mathbf{V}_j^{UL}\right)^H\right)\left(\mathbf{H}_j^{UL}\right)^H \\ & + \sum_{j=1}^{J} \mathbf{H}_0\left(\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H + \kappa \mathrm{diag}\left(\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H\right)\right)\mathbf{H}_0^H + \sigma_0^2\mathbf{I}_{N_0} \\ & + \beta \sum_{j=1}^{J} \mathrm{diag}\left(\mathbf{H}_0\mathbf{V}_j^{DL}\left(\mathbf{V}_j^{DL}\right)^H\mathbf{H}_0^H\right) + \beta \sum_{j=1}^{K} \mathrm{diag}\left(\mathbf{H}_j^{UL}\mathbf{V}_j^{UL}\left(\mathbf{V}_j^{UL}\right)^H\left(\mathbf{H}_j^{UL}\right)^H\right), \end{aligned} \tag{9}$$

$$\begin{aligned} \mathbf{\Sigma}_j^{DL} \approx & \sum_{i \ne j}^{J} \mathbf{H}_j^{DL}\mathbf{V}_i^{DL}\left(\mathbf{V}_i^{DL}\right)^H\left(\mathbf{H}_j^{DL}\right)^H + \kappa \sum_{i=1}^{J} \mathbf{H}_j^{DL}\mathrm{diag}\left(\mathbf{V}_i^{DL}\left(\mathbf{V}_i^{DL}\right)^H\right)\left(\mathbf{H}_j^{DL}\right)^H \\ & + \sum_{k=1}^{K} \mathbf{H}_{jk}^{DU}\left(\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H + \kappa \mathrm{diag}\left(\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H\right)\right)\left(\mathbf{H}_{jk}^{DU}\right)^H + \sigma_j^2\mathbf{I}_{N_j} \\ & + \beta \sum_{k=1}^{K} \mathrm{diag}\left(\mathbf{H}_{jk}^{DU}\mathbf{V}_k^{UL}\left(\mathbf{V}_k^{UL}\right)^H\left(\mathbf{H}_{jk}^{DU}\right)^H\right) + \beta \sum_{i=1}^{J} \mathrm{diag}\left(\mathbf{H}_j^{DL}\mathbf{V}_i^{DL}\left(\mathbf{V}_i^{DL}\right)^H\left(\mathbf{H}_j^{DL}\right)^H\right). \end{aligned} \tag{10}$$

[2]In this paper we consider a system with linear and single-stream decoding. Other decoding possibilities, e.g., successive interference reduction schemes based on single user decoding, can be also considered at the expense of higher complexity, see, e.g., [39], [40].

[3]Note that $\mathbf{\Sigma}_k^{UL}$ and $\mathbf{\Sigma}_j^{DL}$ are approximated under $\kappa \ll 1$ and $\beta \ll 1$, and thus the terms including $\kappa\beta$, i.e., the multiplication of κ and β have been ignored. This is a practical assumption [4], [7]. Note that although the terms κ and β are much smaller than 1, when they are applied on a strong channel alone, i.e., SI channel, they are no longer negligible [7].

[4]In practice, since the BS knows the codeword $\mathbf{x}_0$ (its own transmitted signal), and the SI channel $\mathbf{H}_0$, the term $\mathbf{H}_0\mathbf{x}_0$ can be canceled out in (2). In the following sections, we will keep this term merely to be able to use the simplification in the next subsection. However, in the performance simulations, this term will not be considered.

$i \in \mathcal{S} \triangleq \mathcal{S}^{UL} \cup \mathcal{S}^{DL}$ can be written as

$$\begin{aligned}&\mathbf{MSE}_i\\&= \left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)^H + \mathbf{U}_i^H\mathbf{\Sigma}_i\mathbf{U}_i,\end{aligned} \quad (13)$$

where

$$\begin{aligned}\mathbf{\Sigma}_i = &\sum_{j\in\mathcal{S},j\neq i}\mathbf{H}_{ij}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ij}^H + \kappa\sum_{j\in\mathcal{S}}\mathbf{H}_{ij}\text{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\mathbf{H}_{ij}^H\\&+\beta\sum_{j\in\mathcal{S}}\text{diag}\left(\mathbf{H}_{ij}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ij}^H\right) + \sigma_i^2\mathbf{I}_{\tilde{N}_i},\end{aligned} \quad (14)$$

and the interference power at the l-th PU, I_l^{PU} in (11) can be rewritten as

$$I_l^{PU} = \sum_{j\in\mathcal{S}}\text{tr}\left\{\mathbf{G}_{lj}\left(\mathbf{V}_j\mathbf{V}_j^H + \kappa\text{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\right)\mathbf{G}_{lj}^H\right\}. \quad (15)$$

With the simplified notations, the non-convex optimization problem (12) can be equivalently expressed as

$$\begin{aligned}\min_{\mathbf{V},\mathbf{U}}\quad &\sum_{i\in\mathcal{S}}\text{tr}\left\{\mathbf{MSE}_i\right\} && (16a)\\ \text{s.t.}\quad &\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_i,\ i\in\mathcal{S}^{UL}, && (16b)\\ &\sum_{i\in\mathcal{S}^{DL}}\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_0, && (16c)\\ &I_l^{PU} \le \lambda_l,\ l=1,\ldots,L, && (16d)\end{aligned}$$

where (16d) represents the per PU received interference power constraint. Note that underlay CR systems enable SUs to transmit with overlapping spectrum with PUs as long as the QoS of PUs is not degraded. This is managed by, e.g., introducing some interference constraints that impose upper bounds on the total aggregate interference induced by all SUs to each PU. The choice of this upper bound (or threshold) is a complex and open regulatory issue, which can be the result of a negotiation or opportunistic-based procedure between PUs (or regulatory agencies) and SUs [35]. Both deterministic and probabilistic interference constraints have been suggested in the literature [16], [17]. In this paper, we will consider deterministic interference constraints as assumed in [35], [36], [37]. Particularly, we assume that the PU imposing the interference constraint, has already computed its maximum tolerable interference threshold.

B. Imperfect CSI Model

In practice, perfect CSI of the PUs may not be available at the BS, because of the loose cooperation between PUs and SUs. Therefore, in this paper, we assume that the secondary BS has an imperfect knowledge of the secondary and primary networks. The channel uncertainties are modeled by the worst-case (norm-bounded error) model [19]-[20]:

$$\mathbf{H}_{ij} \in \mathcal{H}_{ij} = \left\{\tilde{\mathbf{H}}_{ij} + \mathbf{\Delta_i} : \|\mathbf{\Delta_i}\|_F \le \delta_i,\ j\in\mathcal{S}\right\}, \quad (17)$$

$$\mathbf{G}_{lj} \in \mathcal{G}_{lj} = \left\{\tilde{\mathbf{G}}_{lj} + \mathbf{\Lambda_l} : \|\mathbf{\Lambda_l}\|_F \le \theta_l,\ j\in\mathcal{S}\right\}, \quad (18)$$

where $\tilde{\mathbf{H}}_{ij}$ and $\tilde{\mathbf{G}}_{lj}$ denote the nominal value of the CSI, $\mathbf{\Delta_i}$ and $\mathbf{\Lambda_l}$ denote the channel error matrix, and δ_i and θ_l denote the uncertainty bounds.

Under channel uncertainties, the problem (16) can be expressed as

$$\begin{aligned}\min_{\mathbf{V},\mathbf{U}}\ \max_{\forall\mathbf{H}_{ij}\in\mathcal{H}_{ij}}\quad &\sum_{i\in\mathcal{S}}\text{tr}\left\{\mathbf{MSE}_i\right\} && (19a)\\ \text{s.t.}\quad &\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_i,\ i\in\mathcal{S}^{UL}, && (19b)\\ &\sum_{i\in\mathcal{S}^{DL}}\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_0, && (19c)\\ &I_l^{PU} \le \lambda_l,\ \forall\mathbf{G}_{lj}\in\mathcal{G}_{lj},\ l=1,\ldots,L. && (19d)\end{aligned}$$

TABLE I: Sum-MSE Minimization using SDP Algorithm

1) Set the iteration number $n=0$ and initialize $\mathbf{V}^{[n]}$.
2) $n \leftarrow n+1$. Update $\mathbf{U}_i^{[n]},\ i\in\mathcal{S}$ by solving the convex SDP problem (21) under fixed $\mathbf{V}^{[n-1]}$.
3) Update $\mathbf{V}_i^{[n]},\ i\in\mathcal{S}$ by solving the convex SDP (21) under fixed $\mathbf{U}^{[n]}$.
4) Repeat steps 2 and 3 until convergence.

Due to the constraint (19d), the problem (19) is a semi-infinite program [24, Ch. 3], and we will derive an equivalent constraint in LMI form in Section III, so that the problem (19) will turn into an equivalent SDP, which can be efficiently solved by standard interior point methods. Then, in an attempt to further reduce the complexity of the SDP algorithm, in Section IV, we will develop a cutting-set based algorithm to solve the non-convex problem (19).

III. SDP METHOD

Since the problem (19) is an intractable semi-infinite optimization problem [41], in the following, we will turn it into a tractable form. Using epigraph form [42, Sec. 4.1.3] and introducing slack variables τ_i, the min-max problem (19) can be equivalently rewritten as the following minimization problem:

$$\begin{aligned}\min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau}}\quad &\sum_{i\in\mathcal{S}}\tau_i && (20a)\\ \text{s.t.}\quad &\text{tr}\left\{\mathbf{MSE}_i\right\} \le \tau_i,\ \forall\mathbf{H}_{ij}\in\mathcal{H}_{ij},\ i\in\mathcal{S}, && (20b)\\ &\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_i,\ i\in\mathcal{S}^{UL}, && (20c)\\ &\sum_{i\in\mathcal{S}^{DL}}\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_0, && (20d)\\ &I_l^{PU} \le \lambda_l,\ \forall\mathbf{G}_{lj}\in\mathcal{G}_{lj},\ l=1,\ldots,L, && (20e)\end{aligned}$$

where $\boldsymbol{\tau}$ is a stacked vector composed of $\tau_i,\ i\in\mathcal{S}$.

The problem (20) can be formulated as a standard SDP, which is defined as minimizing a linear objective under LMI constraints. LMI is a matrix constraint in the form of $\mathbf{A}(\mathbf{x}) \succeq \mathbf{0}$, where the matrix $\mathbf{A}$ depends linearly on $\mathbf{x}$. Thanks to this formulation, many well known algorithms for solving SDPs, e.g., interior point methods [42] can be exploited to solve the optimization problem efficiently in polynomial time. The SDP formulation, equivalent to the problem (20) is expressed as below, the lengthy proof of which is relegated to Appendix A[5].

$$\begin{aligned}\min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau},\epsilon_i\ge0,\eta_l\ge0}\quad &\sum_{i\in\mathcal{S}}\tau_i && (21a)\\ \text{s.t.}\quad &\begin{bmatrix}\tau_i-\epsilon_i & \tilde{\boldsymbol{\mu}}_i^H & \mathbf{0}_{1\times\tilde{N}_i\tilde{M}}\\ \tilde{\boldsymbol{\mu}}_i & \mathbf{I}_{A_i} & -\delta_i\mathbf{D}_{\Delta_i}\\ \mathbf{0}_{\tilde{N}_i\tilde{M}\times1} & -\delta_i\mathbf{D}_{\Delta_i}^H & \epsilon_i\mathbf{I}_{\tilde{N}_i\tilde{M}}\end{bmatrix} \succeq 0,\ i\in\mathcal{S}, && (21b)\\ &\|vec\left(\mathbf{V}_i\right)\|_2^2 \le P_i,\ i\in\mathcal{S}^{UL}, && (21c)\\ &\|\lfloor vec\left(\mathbf{V}_i\right)\rfloor_{i\in\mathcal{S}^{DL}}\|_2^2 \le P_0, && (21d)\\ &\begin{bmatrix}\lambda_l-\eta_l & \tilde{\boldsymbol{\iota}}_l^H & \mathbf{0}_{1\times T_l\tilde{M}}\\ \tilde{\boldsymbol{\iota}}_l & \mathbf{I}_{B_l} & -\theta_l\mathbf{E}_{\Lambda_l}\\ \mathbf{0}_{T_l\tilde{M}\times1} & -\theta_l\mathbf{E}_{\Lambda_l}^H & \eta_l\mathbf{I}_{T_l\tilde{M}}\end{bmatrix} \succeq 0,\ l=1,\ldots,L. && (21e)\end{aligned}$$

The variables A_i, B_l, $\tilde{\boldsymbol{\mu}}_i$, $\mathbf{D}_{\Delta_i}$, $\tilde{\boldsymbol{\iota}}_l$ and $\mathbf{E}_{\Lambda_l}$ are defined as:

[5]To simplify the presentation, from now on we will assume the number of transmit antennas at the BS is equal to number of transmit antennas at the UL users, i.e., $\tilde{M} = M_0 = M_i,\ i\in\mathcal{S}^{UL}$.

$$A_i = d_i \left(\sum_{j\in\mathcal{S}} \left(d_j + \tilde{M}_j \right) + \tilde{N}_i \right) + \tilde{N}_i \sum_{j\in\mathcal{S}} d_j, \quad (22)$$

$$B_l = T_l \sum_{j\in\mathcal{S}} \left(d_j + \tilde{M}_j \right), \quad (23)$$

$$\tilde{\boldsymbol{\mu}}_i = \begin{bmatrix} \left(\mathbf{V}_i^T \otimes \mathbf{U}_i^H\right) vec\left(\tilde{\mathbf{H}}_{ii}\right) - vec\left(\mathbf{I}_{d_i}\right) \\ \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{U}_i^H\right) vec\left(\tilde{\mathbf{H}}_{ij}\right) \right\rfloor_{j\in\mathcal{S}, j\neq i} \\ \left\lfloor \left\lfloor \sqrt{\kappa} \left((\boldsymbol{\Gamma}_\ell \mathbf{V}_j)^T \otimes \mathbf{U}_i^H \right) vec\left(\tilde{\mathbf{H}}_{ij}\right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \\ \left\lfloor \left\lfloor \sqrt{\beta} \left(\mathbf{V}_j^T \otimes (\mathbf{U}_i^H \boldsymbol{\Gamma}_\ell) \right) vec\left(\tilde{\mathbf{H}}_{ij}\right) \right\rfloor_{\ell\in\mathcal{D}_i^{(R)}} \right\rfloor_{j\in\mathcal{S}} \\ \sigma_i vec\left(\mathbf{U}_i\right) \end{bmatrix}, \quad (24)$$

$$\mathbf{D}_{\Delta_i} = \begin{bmatrix} \left(\mathbf{V}_i^T \otimes \mathbf{U}_i^H\right) \\ \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{U}_i^H\right) \right\rfloor_{j\in\mathcal{S}, j\neq i} \\ \left\lfloor \left\lfloor \sqrt{\kappa} \left((\boldsymbol{\Gamma}_\ell \mathbf{V}_j)^T \otimes \mathbf{U}_i^H \right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \\ \left\lfloor \left\lfloor \sqrt{\beta} \left(\mathbf{V}_j^T \otimes (\mathbf{U}_i^H \boldsymbol{\Gamma}_\ell) \right) \right\rfloor_{\ell\in\mathcal{D}_i^{(R)}} \right\rfloor_{j\in\mathcal{S}} \\ \mathbf{0}_{d_i \tilde{N}_i \times \tilde{N}_i \tilde{M}} \end{bmatrix}, \quad (25)$$

$$\tilde{\boldsymbol{\iota}}_l = \begin{bmatrix} \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{I}_{T_l}\right) vec\left(\tilde{\mathbf{G}}_{lj}\right) \right\rfloor_{j\in\mathcal{S}} \\ \sqrt{\kappa} \left\lfloor \left\lfloor \left((\boldsymbol{\Gamma}_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_{T_l} \right) vec\left(\tilde{\mathbf{G}}_{lj}\right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \end{bmatrix}, \quad (26)$$

$$\mathbf{E}_{\Lambda_l} = \begin{bmatrix} \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{I}_{T_l}\right) \right\rfloor_{j\in\mathcal{S}} \\ \sqrt{\kappa} \left\lfloor \left\lfloor \left((\boldsymbol{\Gamma}_\ell \mathbf{V}_j)^T \otimes \mathbf{I}_{T_l} \right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \end{bmatrix}. \quad (27)$$

As it can be observed from (21b), the problem (21) does not hold a jointly convex structure over the optimization variables. Nevertheless, it is a separately convex optimization problem over the transmit beamforming matrices $\mathbf{V}$, and the receiving beamforming matrices $\mathbf{U}$, once the other variables are fixed. This facilitates an alternating optimization algorithm where in each iteration the solution to (21) is calculated, as a convex optimization problem, assuming an alternatively fixed $\mathbf{V}$ or $\mathbf{U}$. The described optimization iterations continue until convergence or a maximum number of iterations is reached. Please see Table I for a detailed algorithm description.

The proposed SDP method decreases the sum MSE monotonically at each iteration of the alternating algorithm. And since MSE is bounded below by zero, it is quite obvious that the proposed sum-MSE minimization algorithm converges [43], [44]. However, the resulting converging point does not necessarily satisfy the first order optimality conditions and since the sum-MSE minimization problem is not jointly convex over transmit and receive filters, the proposed algorithm does not necessarily converge to an optimum point of the original problem. As a result, it is important to select good initialization points to achieve a suboptimal solution with a good performance. Due to the problem complexity, the utilization of such methods with similar convergence result is very common [45]. The performance of the proposed SDP-based algorithm is analyzed via numerical simulations in Section VI.

A. Computational Complexity

In this subsection, we discuss the computational complexity of the proposed SDP algorithm in Table I. The number of arithmetic operations required to solve a standard real-valued SDP problem

$$\min_{\mathbf{x}\in\mathcal{R}^n} \quad \mathbf{c}^T\mathbf{x} \quad (28a)$$

$$\text{s.t.} \quad \mathbf{A}_0 + \sum\nolimits_{i=1}^{n} x_i \mathbf{A}_i \succeq \mathbf{0}, \quad (28b)$$

$$\|\mathbf{x}\|_2 \leq R, \quad (28c)$$

where $\mathbf{A}_i$ denotes the symmetric block-diagonal matrices with P diagonal blocks of size $a_l \times a_l,\ l = 1, \ldots, P$, is upper-bounded by [41]

$$\mathcal{O}(1) \left(1 + \sum_{l=1}^{P} a_l\right)^{1/2} n \left(n^2 + n\sum_{l=1}^{P} a_l^2 + \sum_{l=1}^{P} a_l^3\right). \quad (29)$$

Since the proposed algorithm in Table I solves a SDP problem in Step 2 and Step 3, the number of arithmetic operations required to compute optimal $\mathbf{V}_i$ and $\mathbf{U}_i$ is calculated from (29) as follows. In computing $\mathbf{V}_i$, the number of diagonal blocks P is equal to $|\mathcal{S}| + \left|\mathcal{S}^{UL}\right| + L + 1$. For the MSE constraint of each user, the dimension of blocks are $a_i = A_i + \tilde{N}_i\tilde{M} + 1,\ i \in \mathcal{S}$. For the UL SU power constraint, the dimension of the blocks are $a_i = \tilde{M} d_i^{UL} + 1,\ i \in \mathcal{S}^{UL}$. For the BS power constraint, the dimension of the block is $a_i = \tilde{M} \sum_{i\in\mathcal{S}^{DL}} d_i^{DL} + 1$, and for the PU interference constraint, the dimension of the blocks are $a_l = B_l + T_l\tilde{M} + 1,\ l, \ldots, L$. The size of unknown variables is $n = \sum_{i\in\mathcal{S}} 2\tilde{M} d_i + 2|\mathcal{S}| + L$, where the first term is due to the real and image parts of $\mathbf{V}_i$ and the other terms are due to the additional slack variables. The calculation of the number of arithmetic operations required for $\mathbf{U}_i$ can be carried out similarly. The computational complexity to solve the sum-MSE minimization problem using SDP method are given in Table II. It is easy to see that as the number of users and transmit/receive antennas increase, the complexity of the SDP algorithm is very high.

TABLE II: Complexity of SDP method

	Number of variables (n)	**Dimension of blocks (a_i)**
$\mathbf{V}$	$\sum_{i\in\mathcal{S}} 2\tilde{M} d_i + 2\|\mathcal{S}\| + L$	$a_i = A_i + \tilde{N}_i\tilde{M} + 1,\ i \in \mathcal{S}$ $a_i = \tilde{M} d_i^{UL} + 1,\ i \in \mathcal{S}^{UL}$ $a_i = \tilde{M} \sum_{i\in\mathcal{S}^{DL}} d_i^{DL} + 1$ $a_l = B_l + T_l\tilde{M} + 1,\ l, \ldots, L$
$\mathbf{U}_i$	$2\tilde{N}_i d_i + 2$	$a_i = A_i + \tilde{N}_i\tilde{M} + 1,\ i \in \mathcal{S}$

B. CSI Acquisition

We assume that the secondary BS has the knowledge of the nominal channels and the radius of uncertainty regions. We undertake a centralized approach where the secondary BS collects all channel matrices, computes the beamforming matrices based on the imperfect CSI, and then distributes them to the SUs. The estimation of CSI matrices in the secondary network follows a similar strategy to that of traditional systems, as the secondary nodes cooperate with the secondary BS. This is performed via the exchange of the training sequences and feedback, and the application of usual CSI estimation methods [46]. On the other hand, it is more challenging to obtain an accurate estimate for the CSI between the secondary and primary networks, as the primary network is usually not willing to cooperate with the secondary network. In this regard, few methods have been suggested to combat this problem. Firstly, in case the primary system adopts the TDD scheme, the secondary network can obtain the CSI to the primary nodes by taking advantage of the channel reciprocity, and overhearing the transmissions from the primary network [46]-[50]. Secondly, a partial CSI can be

obtained via blind environmental learning [51], [52]. Third, an estimate of CSI can be obtained via the realization of a *band manager* with the ability to exchange the CSI between the secondary and primary networks [47], [48], [53], and finally, if possible, the primary and secondary networks can cooperate for the exchange of channel estimates [46]. Since we have loose cooperation between primary and secondary networks, the channel estimates will not be perfect in practice. Therefore, we model the imperfect channel estimates with a norm-bounded channel uncertainty model. Note that in SDP-based algorithm, the transmit and receive filters are designed at a central scheduler design with the help of CSI feedbacks at the secondary network [49].

IV. Robust Transceiver Design Based on Cutting-Set Method

Since the SDP algorithm has a high computational complexity, in this section, we propose an efficient with a low complexity algorithm, the cutting-set method [25], which tackles the channel uncertainties by separating worst-case optimization from the robust transceiver design problem. In particular, the optimization problem is solved through alternating between two steps, i.e., transceiver design and worst-case channel determination steps. In the first step (transceiver design), the optimal beamforming matrices are computed under the assumption that the errors belong to a certain known uncertainty region (fixed set of CSI), whereas the second step (worst-case channel determination) computes the worst-case channel error matrices in the uncertainty region that maximize the constraints under the assumption that transceiver beamforming matrices computed in the first step are fixed. In the following, we give both steps of the cutting-set algorithm in detail to solve the robust MSE-based optimization problem.

A. Transceiver design under fixed channels

In this step, a version of the semi-infinite problem is solved over finite subsets of the uncertainty regions. Assuming that the worst-case channels are given (fixed CSI), the optimal $\mathbf{V}_i$ and $\mathbf{U}_i$ are computed through solving the following problem:

$$\min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau}} \quad \sum_{i\in\mathcal{S}} \tau_i \tag{30a}$$

$$\text{s.t.} \quad \|\boldsymbol{\mu}_i\|_2^2 \le \tau_i,\ i\in\mathcal{S}, \tag{30b}$$

$$\|vec\left(\mathbf{V}_i\right)\|_2^2 \le P_i,\ i\in\mathcal{S}^{UL}, \tag{30c}$$

$$\|\lfloor vec\left(\mathbf{V}_i\right)\rfloor_{i\in\mathcal{S}^{DL}}\|_2^2 \le P_0, \tag{30d}$$

$$\|\boldsymbol{\iota}_l\|_2^2 \le \lambda_l,\ l=1,\ldots,L, \tag{30e}$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\iota}_l$ are defined in (54) and (55), respectively. Note that this problem is similar to the optimization problem (56) (without CSI errors) given in Appendix A, and with straightforward manipulations, it is easy to show that the SDP formulation without CSI errors would reduce to the SOCP formulation. Hence, we can reformulate the problem (30) as a SOCP problem under fixed $\mathbf{V}$ or fixed $\mathbf{U}$ [12].

B. Worst-case Channel Determination for Given Transceivers

In the second step, worst-case analysis is carried out where channels that violate the constraints are determined and appended to the finite uncertainty subsets. For fixed transceiver beamforming matrices computed in the first step, the worst-case channels, which maximize the MSE and PU interference constraints given in (30b) and (30e), respectively, are computed in the bounded uncertainty regions. Note that under fixed transceiver beamforming matrices, the MSE and PU interference constraints are independent of each other with respect to the CSI errors. Under a fixed transceiver design, the worst-case channels that maximize the MSE of the i-th user is computed through solving the following problem

$$\max_{\boldsymbol{\Delta_i}} \quad \text{tr}\left\{\mathbf{MSE}_i\right\} \tag{31a}$$

$$\text{s.t.} \quad \|\boldsymbol{\Delta_i}\|_F \le \delta_i, \tag{31b}$$

where $\boldsymbol{\Delta_i}$ is the channel estimation error defined in (17).

Since the function (31a) is non-convex, the problem is intractable, and thus to simplify the computation, we adopt a first order approximation by neglecting all the second and higher terms of CSI errors in (13), similar to [54], [55]. The approximation is expressed as

$$\text{tr}\left\{\mathbf{MSE}_i\right\} \overset{(a)}{\approx} \text{tr}\left\{\tilde{\mathbf{MSE}}_i\right\} + 2\Re\left\{vec^H\left(\mathbf{B}_i^H\right) vec\left(\boldsymbol{\Delta_i}\right)\right\} \tag{32}$$

where $\text{tr}\left\{\tilde{\mathbf{MSE}}_i\right\}$ is obtained by setting all CSI errors in (13) to zero, (a) is obtained by using the identity $\text{tr}\left\{\mathbf{AB}\right\} = vec^H\left(\mathbf{A}^H\right) vec\left(\mathbf{B}\right)$, and $\mathbf{B}_i$ is expressed as

$$\begin{aligned}\mathbf{B}_i = \mathbf{V}_i\left(\mathbf{U}_i^H\tilde{\mathbf{H}}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)^H\mathbf{U}_i^H + \sum_{j\in\mathcal{S},j\neq i}\mathbf{V}_j\mathbf{V}_j^H\tilde{\mathbf{H}}_{ij}^H\mathbf{U}_i\mathbf{U}_i^H \\ + \kappa\sum_{j\in\mathcal{S}}\text{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\tilde{\mathbf{H}}_{ij}^H\mathbf{U}_i\mathbf{U}_i^H \\ + \beta\sum_{j\in\mathcal{S}}\mathbf{V}_j\mathbf{V}_j^H\tilde{\mathbf{H}}_{ij}^H\text{diag}\left(\mathbf{U}_i\mathbf{U}_i^H\right).\end{aligned} \tag{33}$$

Using Cauchy-Schwarz inequality in the approximate MSE expression in (32), the worst-case CSI errors corresponding to the MSE constraints are computed as

$$\boldsymbol{\Delta_i} = \frac{\delta_i}{\|vec\left(\mathbf{B}_i\right)\|_2}\mathbf{B}_i^H. \tag{34}$$

After computing the worst-case CSI errors, their corresponding worst-case channels are written as $\mathbf{H}_{ij} = \tilde{\mathbf{H}}_{ij} + \boldsymbol{\Delta_i},\ j\in\mathcal{S}$. Note that since the worst-case CSI error in (34) is computed through an approximation in (32), it is possible for the MSE constraints in (30b) to be violated even if the corresponding worst-case channels are in the given uncertainty regions. However, when the CSI error is small, this violation is negligible, since the contribution of the second and higher order terms of CSI errors on the performance is also negligible [54], [55].

Next, we compute the worst-case CSI errors corresponding to the PU interference (20e). In particular, under fixed transceiver design, the worst-case channels that maximize the interference power at the l-th PU is computed by solving the following problem

$$\max_{\boldsymbol{\Lambda_l}} \quad I_l^{PU} \tag{35a}$$

$$\text{s.t.} \quad \|\boldsymbol{\Lambda_l}\|_F \le \theta_l, \tag{35b}$$

which is, again, difficult to solve. Here, $\boldsymbol{\Lambda_l}$ is the channel estimation error defined in (18). Similar to (32), an approximation for (15) involving only the first-order errors is adopted, and the approximation is expressed as

$$I_l^{PU} \approx \tilde{I}_l^{PU} + 2\Re\left\{vec^H\left(\mathbf{C}_l^H\right)\ vec\left(\boldsymbol{\Lambda_l}\right)\right\}, \tag{36}$$

where $\tilde{I}_l^{PU}$ is obtained by setting all CSI errors in (15) to zero, and $\mathbf{C}_l$ is defined as

$$\mathbf{C}_l = \sum_{j\in\mathcal{S}}\left(\mathbf{V}_j\mathbf{V}_j^H + \kappa\text{diag}\left(\mathbf{V}_j\mathbf{V}_j^H\right)\right)\tilde{\mathbf{G}}_{lj}^H. \tag{37}$$

Using Cauchy-Schwarz inequality in the approximate expression (36), the worst-case errors corresponding to the l-th PU interference are obtained as

$$\mathbf{\Lambda_l} = \frac{\theta_l}{\|vec\,(\mathbf{C}_l)\|_2}\mathbf{C}_l^H. \tag{38}$$

The corresponding worst-case channels can be expressed as $\mathbf{G}_{lj} = \tilde{\mathbf{G}}_{lj} + \mathbf{\Lambda_l},\ j \in \mathcal{S}$.

C. Iterative Algorithm for the Robust Design

The proposed cutting-set algorithm to solve the robust MSE-based problem involves a two-step algorithm alternating over transceiver optimization and worst-case analysis steps, where at each step a convex optimization problem is solved as described in the previous two subsections. The algorithm starts with the set of channel matrices $\mathcal{H}$, which initially contains only the imperfect CSI $\tilde{\mathbf{H}}_{ij},\ \{i,j\} \in \mathcal{S}$ and $\tilde{\mathbf{G}}_{lj},\ j \in \mathcal{S},\ l = 1,\ldots,L$. In the first step, the problem (30) is solved under all the channels in $\mathcal{H}$ (the constraints involve all the channel elements in the set $\mathcal{H}$).

The second step involves computing the worst-case channels through the problems (31) and (35) under the transceivers obtained in the previous step. After the worst-case analysis, if the corresponding channels for all links, i.e., $\tilde{\mathbf{H}}_{ij} + \mathbf{\Delta_i}$ and $\tilde{\mathbf{G}}_{lj} + \mathbf{\Lambda_l}$ violate the constraints in (30b) and/or (30e) under the transceivers obtained in the first step, these worst-channels are appended to the set $\mathcal{H}$. The alternating two-step algorithm runs until no violating channel is produced, i.e., the maximum constraint violation is below a specified threshold [54], [55], [56], [57]. Note that during the worst-case channel determination step, the set $\mathcal{H}$ may be expanded (or remain the same) depending on the constraint violations. During the minimization step (first step), the precoder and receive beamforming matrices are computed to meet MSE and PU interference constraints for increasing number of worst-case channels in $\mathcal{H}$ (MSE and PU interference constraints must include all the channels in $\mathcal{H}$) resulting in increased robustness. Particularly, as the size of the set $\mathcal{H}$ increases, the number of effective constraints in the transceiver design problem increases.

As shown in [25], the two-step alternating algorithm converges to an optimal solution of the original problem when the optimization and worst-case analysis steps are solved to global optimality in each iteration. However, the optimization step here involves solving the non-convex problem in (30) and furthermore due to the MSE and interference approximations in (32) and (36), respectively used to convexify the optimization problem in the worst-case analysis, the proposed iterative algorithm is not guaranteed to lead to the robust optimal solution, and generally converges to a suboptimal solution. Since the optimization problem is intractable, the convergence of the proposed cutting-set algorithm is not shown analytically but only demonstrated empirically, similar to the prior related works [54], [55], [56], [57] . However, our simulations show that the proposed design is robust to the CSI errors, i.e., it can still provide a significant gain over the non-robust method in the presence of CSI errors. Moreover, using warm-start [25] techniques in the outer iterations, i.e., use the previously computed precoder and receive filter to initialize the solution of the current outer iteration reduces the overall effort for our proposed scheme to converge. The steps of the proposed cutting-set method is shown in Table III.

TABLE III: Sum-MSE Minimization using Cutting-Set Method

1) Set the iteration number $n = 0$ and initialize $\mathcal{H}^{[n]}$.
2) Update $\mathbf{U}_i^{[n]}$ and $\mathbf{V}_i^{[n]}$ by solving the problem (30) with the given set $\mathcal{H}^{[n]}$ utilizing a similar iterative algorithm given in Table I.
3) Compute the worst-case channels, denoted as $\tilde{\mathcal{H}}_{[n]}$ using (34) and (38).
4) Find the violating channels and append them to the set, i.e., $\mathcal{H}^{[n+1]} = \left\{\tilde{\mathcal{H}}_{[n]}, \mathcal{H}^{[n]}\right\}$.
5) If the termination criterion is satisfied, then end. Otherwise, set $n \leftarrow n+1$ and go to Step 2.

TABLE IV: Complexity of cutting-set method

	Number of variables (n)	**Dimension of blocks** (a_i)
$\mathbf{V}$	$\sum_{i\in\mathcal{S}} 2\tilde{M}d_i + |\mathcal{S}|$	$a_i = A_i - \tilde{N}_i d_i,\ i \in \mathcal{S}$ $a_i = \tilde{M}d_i^{UL},\ i \in \mathcal{S}^{UL}$ $a_i = \tilde{M}\sum_{i\in\mathcal{S}^{DL}} d_i^{DL}$ $a_l = B_l,\ l,\ldots,L$
$\mathbf{U}_i$	$2\tilde{N}_i d_i + 1$	$a_i = A_i,\ i \in \mathcal{S}$

D. Computational Complexity

The main computation complexity of the cutting-set algorithm stems from the SOCP problem (30), since the computation complexity of the worst-case analysis can be neglected. A real-valued SOCP problem can be expressed as

$$\min_{\mathbf{x}\in\mathcal{R}^n} \quad \mathbf{c}^T\mathbf{x} \tag{39a}$$

$$\text{s.t.} \quad \|\mathbf{A}_i\mathbf{x} + \mathbf{b}_i\| \le \mathbf{c}_i^T\mathbf{x} + d_i,\ i = 1,\ldots,P, \tag{39b}$$

$$\|\mathbf{x}\|_2 \le R, \tag{39c}$$

where $\mathbf{b}_i \in \mathcal{R}^{a_i}$. As discussed in [41], the upper bound on the number of arithmetic operations to solve this problem is:

$$\mathcal{O}\,(1)\,(1+P)^{1/2}\,n\left(n^2 + P + \sum_{i=0}^{P} a_i^2\right). \tag{40}$$

In our problem, the number of inequalities $P+1$ equals to $|\mathcal{S}| + \left|\mathcal{S}^{UL}\right| + L + 1$. For the MSE constraint of each user, the dimension of blocks are $a_i = A_i - \tilde{N}_i d_i,\ i \in \mathcal{S}$. For the UL SU power constraint, the dimension of the blocks are $a_i = \tilde{M}d_i^{UL},\ i \in \mathcal{S}^{UL}$. For the BS power constraint, the dimension of the block is $a_i = \tilde{M}\sum_{i\in\mathcal{S}^{DL}} d_i^{DL}$, and for the PU interference constraint, the dimension of the blocks are $a_l = B_l,\ l,\ldots,L$. The size of the unknown variables are $n = \sum_{i\in\mathcal{S}} 2\tilde{M}d_i + |\mathcal{S}|$. By carrying out a similar analysis, the complexity of other subproblems can be computed. The computational complexity for the sum-MSE problem based on cutting-set method is shown in Table IV. Note that compared to the complexity of the SDP method in Table II, the cutting-set method has a lower computational complexity, which will be further compared in the simulations.

V. Extensions

A. Different Utility Functions

In this subsection, it is shown that with minor modifications, the proposed algorithms can also be applied to solve the other robust optimization problems. One problem can be sum-power minimization subject to MSE and interference constraints, which can be formulated as:

$$\min_{\mathbf{V},\mathbf{U}} \quad \sum_{i\in\mathcal{S}} \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \tag{41a}$$

$$\text{s.t.} \quad \text{tr}\left\{\mathbf{MSE}_i\right\} \le \varsigma_i,\ i \in \mathcal{S},\ \forall \mathbf{H}_{ij} \in \mathcal{H}_{ij}, \tag{41b}$$

$$I_l^{PU} \le \lambda_l,\ \forall \mathbf{G}_{lj} \in \mathcal{G}_{lj},\ l = 1,\ldots,L, \tag{41c}$$

where $\varsigma_i,\ i \in \mathcal{S}$ is the MSE constraint of the i-th user. Using epigraph form and introducing slack variables τ_i, the problem

can be written as

$$\min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau}} \quad \sum_{i\in\mathcal{S}} \tau_i \tag{42a}$$

$$\text{s.t.} \quad \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le \tau_i,\ i\in\mathcal{S}, \tag{42b}$$

$$\text{tr}\left\{\mathbf{MSE}_i\right\} \le \varsigma_i,\ i\in\mathcal{S},\ \forall \mathbf{H}_{ij}\in\mathcal{H}_{ij}, \tag{42c}$$

$$I_l^{PU} \le \lambda_l,\ \forall \mathbf{G}_{lj}\in\mathcal{G}_{lj},\ l=1,\ldots,L. \tag{42d}$$

Another optimization problem can be the minimization of maximum per-user MSE subject to power constraints at the secondary network, and interference constraints from the secondary to primary network. Unlike the sum-MSE minimization problem, the min-max per-user MSE problem makes sure that each user has the same MSE, an thus the problem brings fairness among the nodes. The min-max MSE problem is written as:

$$\min_{\mathbf{V},\mathbf{U}} \max_{\forall \mathbf{H}_{ij}\in\mathcal{H}_{ij}, i\in\mathcal{S}} \quad \text{tr}\left\{\mathbf{MSE}_i\right\} \tag{43a}$$

$$\text{s.t.} \quad \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_i,\ i\in\mathcal{S}^{UL}, \tag{43b}$$

$$\sum_{i\in\mathcal{S}^{DL}} \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_0, \tag{43c}$$

$$I_l^{PU} \le \lambda_l,\ \forall \mathbf{G}_{lj}\in\mathcal{G}_{lj},\ \forall l, \tag{43d}$$

which is equivalent to the following problem

$$\min_{\mathbf{V},\mathbf{U},\tau} \quad \tau \tag{44a}$$

$$\text{s.t.} \quad \text{tr}\left\{\mathbf{MSE}_i\right\} \le \tau,\ \forall \mathbf{H}_{ij}\in\mathcal{H}_{ij},\ i\in\mathcal{S}, \tag{44b}$$

$$\text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_i,\ i\in\mathcal{S}^{UL}, \tag{44c}$$

$$\sum_{i\in\mathcal{S}^{DL}} \text{tr}\left\{\mathbf{V}_i\mathbf{V}_i^H\right\} \le P_0, \tag{44d}$$

$$I_l^{PU} \le \lambda_l,\ \forall \mathbf{G}_{lj}\in\mathcal{G}_{lj},\ l=1,\ldots,L. \tag{44e}$$

Since the problems (42) and (44) have similar structures with (20), both problems (42) and (44) can be solved using both SDP and cutting-set methods proposed in Section III and Section IV, respectively.

B. Multiple Uncertainties

So far, we have assumed that each semi-infinite constraint (MSE and interference constraints) includes only one uncertainty variable. Here, we will generalize this to the case of complex-valued quantities with multiple uncertainties in each design constraint. To that end, the channel is modeled as

$$\mathbf{H}_{ij}\in\mathcal{H}_{ij} = \left\{\tilde{\mathbf{H}}_{ij} + \boldsymbol{\Delta}_{ij} : \|\boldsymbol{\Delta}_{ij}\|_F \le \delta_{ij}\right\}, \tag{45}$$

$$\mathbf{G}_{lj}\in\mathcal{G}_{lj} = \left\{\tilde{\mathbf{G}}_{lj} + \boldsymbol{\Lambda}_{lj} : \|\boldsymbol{\Lambda}_{lj}\|_F \le \theta_{lj}\right\}, \tag{46}$$

where $\tilde{\mathbf{H}}_{ij}$, $\tilde{\mathbf{G}}_{lj}$ are defined in (17)-(18), and δ_{ij}, λ_{lj} denote the nominal value of uncertainty bounds. Please note that compared to the model in (17)-(18), which has only one uncertainty variable, the model in (45)-(46) has multiple uncertainties.

To handle the multiple uncertainties, we first rewrite the MSE term in (13) as

$$\boldsymbol{\mu}_i = \begin{bmatrix} \begin{bmatrix} \left(\mathbf{V}_j^T\otimes\mathbf{U}_i^H\right) vec\left(\mathbf{H}_{ij}\right) \\ \left\lfloor\sqrt{\kappa}\left(\left(\boldsymbol{\Gamma}_\ell\mathbf{V}_j\right)^T\otimes\mathbf{U}_i^H\right) vec\left(\mathbf{H}_{ij}\right)\right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \\ \left\lfloor\sqrt{\beta}\left(\mathbf{V}_j^T\otimes\left(\mathbf{U}_i^H\boldsymbol{\Gamma}_\ell\right)\right) vec\left(\mathbf{H}_{ij}\right)\right\rfloor_{\ell\in\mathcal{D}_i^{(R)}} \end{bmatrix}_{j\in\mathcal{S}} \\ \sigma_i vec\left(\mathbf{U}_i\right) \end{bmatrix}$$

$$- \begin{bmatrix} \mathbf{0}_{\left(\sum_{j<i}\left(\tilde{N}_j+\tilde{M}_j+1\right)d_id_j\right)\times 1} \\ vec\left(\mathbf{I}_{d_i}\right) \\ \mathbf{0}_{\left(d_i(\tilde{N}_i-d_j)+\sum_{j\ge i}\left(\tilde{N}_j+\tilde{M}_j+1\right)d_id_j\right)\times 1} \end{bmatrix} \tag{47}$$

$$= \tilde{\boldsymbol{\mu}}_i + \sum_{j\in\mathcal{S}} \mathbf{M}_{ij}\ vec\left(\boldsymbol{\Delta}_{ij}\right), \tag{48}$$

where $\tilde{\boldsymbol{\mu}}_i$ is the same as the first term in (47), and the only difference is that the actual channel matrices $\mathbf{H}_{ij},\ \forall\{i,j\}$ is replaced with the nominal channel matrices $\tilde{\mathbf{H}}_{ij}$, and $\mathbf{M}_{ij}$ is defined as

$$\mathbf{M}_{ij} = \begin{bmatrix} \mathbf{0}_{\left(\sum_{j<i}\left(\tilde{N}_j+\tilde{M}_j+1\right)d_id_j\right)\times\tilde{N}_i\tilde{M}_j} \\ \left(\mathbf{V}_j^T\otimes\mathbf{U}_i^H\right) \\ \left\lfloor\sqrt{\kappa}\left(\left(\boldsymbol{\Gamma}_\ell\mathbf{V}_j\right)^T\otimes\mathbf{U}_i^H\right)\right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \\ \left\lfloor\sqrt{\beta}\left(\mathbf{V}_j^T\otimes\left(\mathbf{U}_i^H\boldsymbol{\Gamma}_\ell\right)\right)\right\rfloor_{\ell\in\mathcal{D}_i^{(R)}} \\ \mathbf{0}_{\left(\tilde{N}_id_i+\sum_{j>i}\left(\tilde{N}_j+\tilde{M}_j+1\right)d_id_j\right)\times\tilde{N}_i\tilde{M}_j} \end{bmatrix}. \tag{49}$$

Using the extended sign-definiteness lemma [44], the MSE constraint $\|\boldsymbol{\mu}_i\|_2^2 \le \tau_i$ in (56b) can be expressed as

$$\begin{bmatrix} \mathbf{A}_i - \sum_{j\in\mathcal{S}}\epsilon_{ij}\mathbf{Q}_{ij}^H\mathbf{Q}_{ij} & -\delta_{i1}\mathbf{P}_{i1}^H & \ldots & -\delta_{i|\mathcal{S}|}\mathbf{P}_{i|\mathcal{S}|}^H \\ \delta_{i1}\mathbf{P}_{i1} & \epsilon_{i1}\mathbf{I}_{\tilde{N}_i\tilde{M}} & \ldots & \mathbf{0}_{\tilde{N}_i\tilde{M}\times\tilde{N}_i\tilde{M}} \\ \vdots & \vdots & \ddots & \vdots \\ -\delta_{i|\mathcal{S}|}\mathbf{P}_{i|\mathcal{S}|} & \mathbf{0}_{\tilde{N}_i\tilde{M}\times\tilde{N}_i\tilde{M}} & \ldots & \epsilon_{i|\mathcal{S}|}\mathbf{I}_{\tilde{N}_i\tilde{M}} \end{bmatrix} \succeq 0,\ i\in\mathcal{S}, \tag{50}$$

$$\epsilon_{ij}\ge 0,\ \forall\{i,j\}, \tag{51}$$

where

$$\mathbf{A}_i = \begin{bmatrix} \tau_i & \tilde{\boldsymbol{\mu}}_i^H \\ \tilde{\boldsymbol{\mu}}_i & \mathbf{I}_{A_i} \end{bmatrix},\ \mathbf{Q}_{ij} = \left[-1,\ \mathbf{0}_{1\times A_i}\right],$$

$$\mathbf{P}_{ij} = \left[\mathbf{0}_{\tilde{N}_i\tilde{M}\times 1},\ \mathbf{M}_{ij}^H\right],$$

and $\epsilon_{ij},\ \forall\{i,j\}$ are non-negative real numbers. The same steps can be independently carried out for the interference power constraint $\|\boldsymbol{\iota}_l\|_2^2 \le \lambda_l$ in (56e). With the LMI forms in hand, the proposed SDP can be applied for the case with the multiple uncertainties.

VI. Simulation Results

In this section, we numerically investigate the sum-MSE minimization problem for an FD MIMO CR cellular system. We start by comparing the sum MSE performance of the two algorithms, proposed in the paper as a function of transmitter/receiver distortion, κ/β and channel uncertainty size, δ/θ. We then analyze the sum-rate performance of the FD system as a function of the number of antennas $\tilde{N}$, transmitter/receiver distortion, κ/β, channel uncertainty size, δ/θ and CCI attenuation factor[6], ν. The tolerance (the difference between MSE of two iterations) of the proposed iterative algorithm is set to 10^{-4}, the maximum number of iterations is set to 50, and the results are averaged over 100 independent channel realizations. Since the optimization problems we are dealing with are non-convex, we need to choose good initialization points to have a suboptimal solution with a good performance. In this paper, we use right singular matrices initialization [58], and have used multiple initializations to improve the performance. For each simulation point, a certain number of initializations were tried and the best result was chosen. We have observed in

[6]It is important to note that while the channel matrices are assumed to be given for each user, it is essential for a practical system to exploit a smart channel assignment algorithm prior to precoder/decoder design. This is particularly essential for an FD setup as the CCI can be reduced by assigning the users with weaker interference paths into the same channel. In order to incorporate the effect of channel assignment into our simulation, we assume an attenuation coefficient, namely ν, on the CCI channels, which represent the degree of isolation among UL and DL users due to channel assignment.

TABLE V: Simulation Parameters

Parameter	Settings
Cell Radius	40m
Carrier Frequency	2GHz
Bandwidth	10MHz
Thermal Noise Density	-174dBm/Hz
Noise Figure	BS: 13dB, User: 9dB
Path Loss (dB) between BS and users (d in km)	$103.8 + 20.9 \log_{10} d$
Path Loss (dB) between users (d in km)	$145.4 + 37.5 \log_{10} d$
Shadowing Standard Deviation	LOS: 3dB, NLOS: 4dB

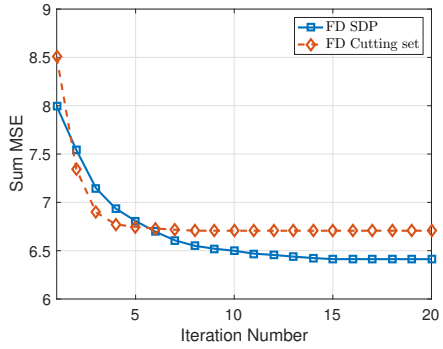


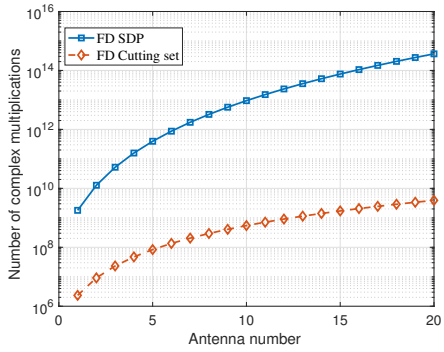
Fig. 2: Convergence behavior of the proposed algorithms.

our simulations that the difference caused by the number of initializations is not very large after 20 initializations. Thus, in the sequel, we have used 20 initializations.

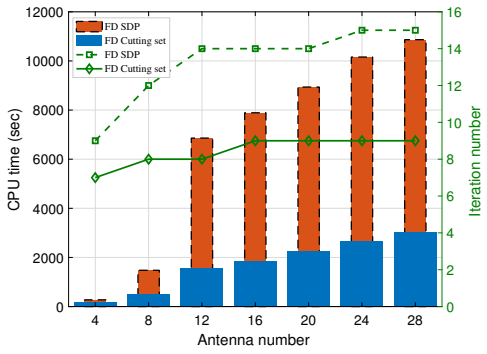
We consider small cell deployments [59] and compare the FD system with the HD system under the 3rd Generation Partnership Project (3GPP) Long-Term Evolution (LTE) specifications. Small cell is considered to be suitable for deployment of FD technology due to its low transmit power, short transmission distances and low mobility [14], [60]. We consider a single hexagonal cell with cell radius $r = 40$m, consisting of a BS in the center with M_0 transmit and N_0 receive antennas. $K = 2$ UL and $J = 2$ DL users equipped with N antennas randomly distributed in the cell[7]. For simplicity, we assume $M_0 = N_0 = N = \tilde{N}$. The CR system has $L = 2$ PUs, with the same maximum allowed interfering power $\lambda_l = 0$dB. The channel between BS and all users are assumed to experience the path loss model for line-of-sight (LOS), and the channel between UL and DL users are assumed to experience the path loss model for non-line-of-sight (NLOS) communications. Detailed simulation parameters are shown in Table V.

The estimated channel gain between the BS to kth UL user is given by $\tilde{\mathbf{H}}_k^{UL} = \sqrt{\kappa_k^{UL}}\hat{\mathbf{H}}_k^{UL}$, where $\hat{\mathbf{H}}_k^{UL}$ denotes the small scale fading following a complex Gaussian distribution with zero mean and unit variance, and $\kappa_k^{UL} = 10^{(-\mathrm{X}/10)}$, $X \in \{\mathrm{LOS}, \mathrm{NLOS}\}$ represents the large scale fading consisting of path loss and shadowing, where LOS and NLOS are calculated from a specific path loss model given in Table V. The channels between BS and DL users, between UL users and DL users, between BS and PUs, and between UL users and PUs are defined similarly. We adopt the Rician model in [3], in which the SI channel is distributed as $\tilde{\mathbf{H}}_0 \sim \mathcal{CN}\left(\sqrt{\frac{K_R}{1+K_R}}\hat{\mathbf{H}}_0, \frac{1}{1+K_R}\mathbf{I}_{N_0} \otimes \mathbf{I}_{M_0}\right)$, where K_R is the Rician factor, and $\hat{\mathbf{H}}_0$ is a deterministic matrix[8]. Unless stated otherwise, we consider, $\tilde{N} = 2$, $\kappa = \beta = -70$dB, $\nu = 0.5$ and $\delta = \theta = 0.1$.

[7]Although the BS has $N_0 + M_0$ antennas in total, we assume that only M_0 (N_0) antennas can be used for transmission (reception) in HD mode. This assumption is similar to [15]. The reason is that in practical systems RF front-ends are scarce resources, since they are much more expensive than antennas. Therefore, we assume that BS only has M_0 transmission front-ends and N_0 receiving front-ends, and do not carry out antenna partitioning.



(a) Complexity comparison.



(b) CPU time and iteration number comparisons.

Fig. 3: Complexity, CPU time and iteration number comparisons of SDP and cutting-set algorithm systems with respect to different number of antennas. In (a), 3 UL, 3 DL, 2 PU equipped with 3 antennas, and $d = 2$ data stream transmission is assumed.

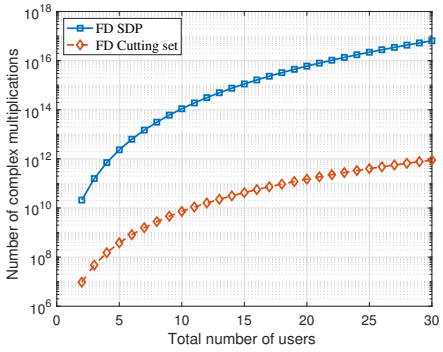
Fig. 2 shows the evolution of the proposed algorithms, i.e., the convergence of the algorithms in Table I and Table III. The monotonic decrease of the sum-MSE can be verified, and is seen that the cutting set algorithm converges more rapidly than SDP.

After establishing the convergence of the two algorithms, we now present a thorough comparison of the SDP and cutting-set methods in terms of computational complexity (complex multiplications) and CPU time (time in secs required for convergence)/iteration number (number of iterations required to converge) in Fig. 3a and Fig. 3b, respectively with respect to different number of antennas. Similarly, in Fig. 4a and Fig. 4b, computational complexity and CPU time/iteration number are plotted with respect to different number of users for the two algorithms, respectively.[9] Note that in Fig. 3b and Fig. 4b, the bar plots represent the CPU time and lines represent the number of iterations. As expected, cutting-set algorithm always has the lowest complexity and requires less computational time than SDP, especially at high number of antennas and users, which is inline with our computational complexity analysis in Table II and IV.

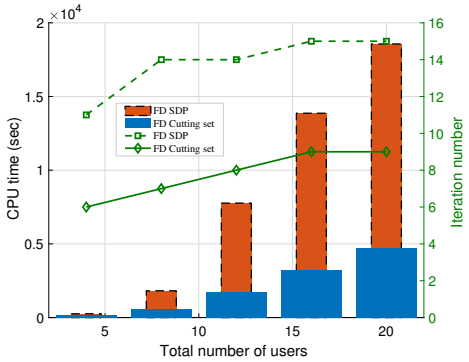
Next, we compare the sum-MSE performance of the proposed SDP and cutting set algorithms under various $\kappa = \beta$

[8]Similar to [14], without loss of generality, we set $K_R = 1$ and $\tilde{\mathbf{H}}_0$ to be the matrix of all ones for all experiments.

[9]For system guidelines we note that, the proposed algorithms are evaluated centrally using MATLAB R20015a on a Linux server with Intel Xeon processor (16 cores, each clocked at 2 GHz) and 31.4 GiB of memory.



(a) Complexity comparison.



(b) CPU time and iteration number comparisons.

Fig. 4: Complexity, CPU time and iteration number comparisons of SDP and cutting-set algorithm systems with respect to different number of users. In (a), 4 transmit/receive antennas, 2 PU equipped with 3 antennas, and $d = 2$ data stream transmission is assumed.

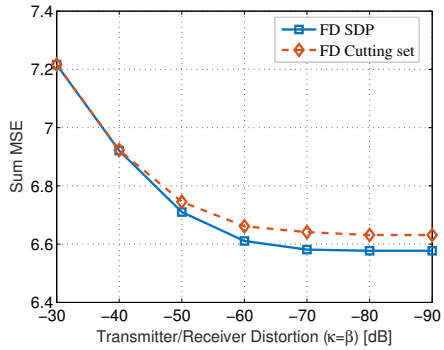


Fig. 5: Sum-MSE comparison of SDP and cutting-set algorithms for an FD system with respect to transmitter/receiver distortion, i.e., κ, β.

values in Fig. 5. From the figure, it can be seen that the performance of the cutting set method is similar to the SDP based one, but with a nominal performance gap. This loss in performance is well compensated from the computational point of view as the cutting set based method provides affordable computational complexity with respect to its SDP counterpart. To further highlight the similarities between the two algorithms in terms of MSE performance, in Fig. 6, we compare the sum-MSE performance of both SDP and cutting set algorithms under different channel uncertainty sizes. When the channel uncertainty size is small, both the algorithms perform similarly, whereas with an increase in $\delta = \theta$ values, the performance gap between the two algorithm increases. The reason is that cutting-set algorithm is derived based on the approximations given in (32) and (36), where the second-order CSI errors are ignored. Therefore, as the channel uncertainty size in-

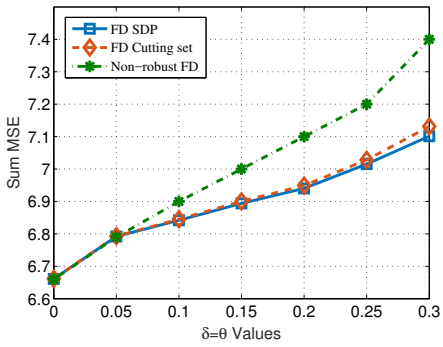


Fig. 6: Sum-MSE comparison of SDP, cutting-set and non-robust algorithms for an FD system with respect to channel uncertainty, $\delta = \theta$. Here, $\kappa = \beta = -40$dB.

creases, the effect of second-order CSI errors become more apparent [55]. In Fig. 6, we also illustrate the performance of the non-robust algorithm. The non-robust beamformer is designed based on the nominal estimated value of the CSI by setting the radius of the uncertainty region to zero. It is shown that the performance of the non-robust design which assumes the estimated CSI as perfect is found to deteriorate as the uncertainty region increases.

Hence, from the above results, we can conclude that the SDP algorithm achieves lower MSE, but with the cost of slow convergence and high CPU time, which is inline with our computational complexity analysis in Table II and IV. On the other hand, the cutting set method can be considered as a good alternative to the SDP method, which offers a decent trade-off between performance and computational complexity.

In our next example, we give the sum-rate comparison of FD and HD systems as a function of $\kappa = \beta$ values for different number of antennas based on the SDP algorithm. We would like to note that hereinafter, we have not included the cutting-set algorithm, since we have observed that both SDP and cutting-set algorithms give very similar performance. The sum-rate of the MIMO FD cellular system assuming a linear single-stream decoding at the receiver can be expressed as

$$I_{sum} = \sum\nolimits_{i\in\mathcal{S}} \sum\nolimits_{k=1}^{d_i} \log_2\left(1 + \mathrm{SINR}_{i_k}\right), \tag{52}$$

where SINR_{i_k} is SINR of the k-th stream of user i defined as

$$\mathrm{SINR}_{i_k} = \frac{\mathbf{u}_{i_k}^H \mathbf{H}_{ii}\mathbf{v}_{i_k}\mathbf{v}_{i_k}^H \mathbf{H}_{ii}^H \mathbf{u}_{i_k}}{\mathbf{u}_{i_k}^H \left(\mathbf{\Sigma}_i + \sum_{j\neq k} \mathbf{H}_{ii}\mathbf{v}_{i_j}\mathbf{v}_{i_j}^H \mathbf{H}_{ii}^H\right)\mathbf{u}_{i_k}}. \tag{53}$$

In (53), $\mathbf{u}_{i_k}$ and $\mathbf{v}_{i_k}$ are the k-th column of $\mathbf{U}_i$, and $\mathbf{V}_i$, respectively. It is seen in Fig. 7 that the performance of the HD system is invariant to transmit/receive distortion κ and β values, as expected and as the SI cancellation capability increases, the sum-rate achieved by the FD system is around 1.6 times more than that of HD. However, at low SI cancellation levels (below around $\kappa = \beta = -55$dB), the distortion is magnified with the increasing number of antennas and the sum-rate of the FD system starts being outperformed by that of HD scheme. Note that transmitter/receiver distortion parameters $\kappa = \beta = -55$dB refers to digital domain SI cancellation. As mentioned in the fourth footnote, the main SI term has been cancelled out in the analog domain, and the residual SI is handled in the digital domain. Similar digital cancellation values have also been reported experimentally in [4].

In Fig. 8, the importance of the smart channel assignment, as a stage prior to the precoder/decoder design is depicted for the SDP algorithm. The CCI attenuation represents the

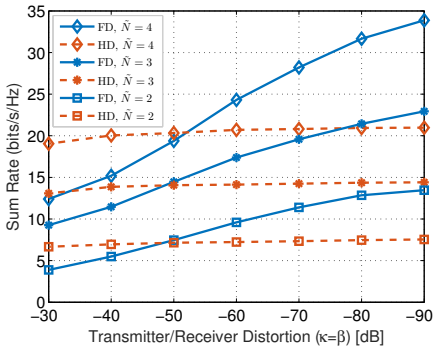


Fig. 7: Sum-rate comparison of FD and HD systems for various antenna numbers under different transmitter/receiver distortion, i.e., κ, β.

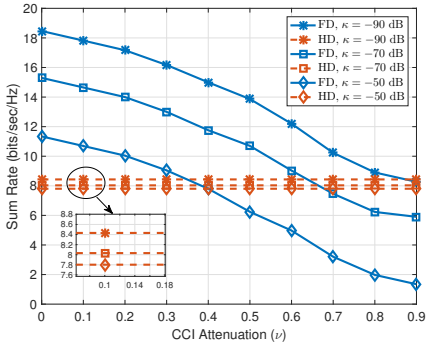


Fig. 8: Sum-rate comparison of FD and HD systems with respect to CCI attenuation factor, i.e., ν.

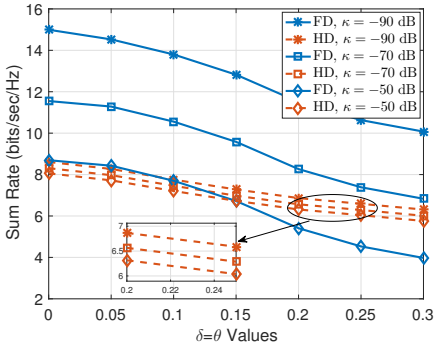


Fig. 9: Sum-rate comparison of FD and HD systems with respect to channel uncertainty, $\delta=\theta$.

provided isolation among the UL and DL users. It is seen that as the suppression level of CCI increases (ν decreases), the FD system starts outperforming the HD system, and thus isolation among the UL and DL users is essential for a successful coexistence of UL and DL users in an FD setup.

In Fig. 9, sum-rate comparison of FD and HD systems under various κ=β values as a function of δ=θ based on the SDP algorithm is shown. It can be seen that the FD results deteriorate more than the HD ones for the same decrease CSI quality. The reason is that when operating in FD mode, there are more channel links between the various nodes than for the corresponding HD systems. Having an increased amount of links with imperfect knowledge result in a sharper rate decrease, thereby stressing the added importance of channel estimation and robust beamformer design for FD systems.

Finally, in Fig. 10 we compare the sum-rate performance of the FD system with respect to path loss. As mentioned previously, the path loss considered in this work is adopted

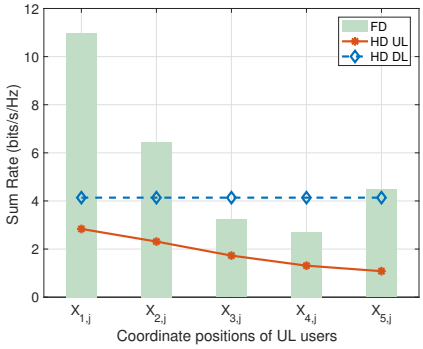


Fig. 10: Sum-rate comparison of FD and HD systems for various coordinate positions of the UL users. The x-axis represents the coordinate positions of the UL users, i.e., $X_{i,j}$, with $j \in \{1,2\}$ and $i=[1,2,3,4,5]$. Here, $||X_{1,j}||_2 < \cdots < ||X_{5,j}||_2$.

from the 3GPP model [59], where it depends on the distance d between the BS and users. Here, d is in meters and is a random variable. In particular, we begin by randomly generating the coordinate positions of the UL and DL users inside the circumference of the cell, while the BS is located at the centre of the cell with coordinates $(0,0)$. The first coordinates for the UL users, $X_{i,j}=(x_{i,j},y_{i,j})$ is generated such that the distance from the BS, $d_{i,j}=||X_{i,j}||_2 \leq 10$m. Now, without altering the coordinates of the DL users, we start displacing the coordinates of the UL users with respect to the BS in steps of 5m until it reaches the cell edge. It can be seen that when the distance of the UL users increases, the sum-rate of the FD system monotonically decreases to reach an optimal point at coordinate position $X_{4,j}$ before increasing again at coordinate position $X_{5,j}$. This can be explained as follows. When the distance of the UL users from the BS increases, the path loss increases, resulting in a decrease of the received signal power at the BS. As a result the residual SI power at the BS overwhelms the received signal power. Furthermore, as seen from Fig. 8, in FD systems the CCI is another important factor that affects FD systems. While displacing the coordinates of the UL users away from the BS, UL users may tend get closer to the DL users, which is when the CCI overwhelms the DL users. As a result, the sum-rate of the FD system falls alarmingly below the sum-rate of the HD system. Here, coordinates $X_{3,j}$ and $X_{4,j}$ are the two worst case scenarios. However, as the UL users move even farther, the CCI on the DL users starts decreasing and the sum-rate of the FD system increases again. At coordinate $X_{5,j}$, the UL of the FD system is almost inoperable and the system is operating only in DL mode. Here, the path loss affecting the UL transmission is so high that the signal power can no longer be differentiated from the residual SI power at the BS. Numerically, average of $X_{5,j}$ is found as $X_{5,1}$=$(26.3764, 23.9888)$ and $X_{5,2}$=$(26.7644, -28.9238)$. At this point the sum-rate of the FD system is comparable to that of the DL of the HD system. Further, it can also be seen from the figure that the UL sum-rate of the HD system decreases monotonically. This is due to the fact that as the distance between the UL users and the BS increases, the path loss increases, which in turn decreases the signal power resulting in reduction of the sum rate in the UL of the HD. However, as the coordinates of the DL users are not altered, the sum-rate for the DL users in HD mode are not affected and remains constant, which can be seen from the figure as well.

VII. Conclusion

We have tackled the sum-MSE minimization problem for an FD MIMO CR cellular system that suffers from the transmit/receive distortions and the imperfect CSI. Two algorithms are proposed to solve this non-convex problem, i) an iterative SDP-based algorithm that computes the transmit and receiving filters jointly, ii) a two-step cutting-set method that alternates between transceiver design and worst-case channel analysis. As simulation results demonstrate, the cutting-set method achieves a similar performance to that of the SDP-based method with a lower computational complexity. Moreover, it has been shown in simulations that replacing standard HD networks with FD ones within this context can indeed increase achievable sum rate for low to intermediate distortion levels.

Appendix A: Problem Reformulation

To solve the optimization problem (20), we first write it in a more compact form for ease of exposition. To that end, we write $\text{tr}\{\mathbf{MSE}_i\}$ and I_l^{PU} in (20) in vector forms. As shown in Appendix B, the vector forms of $\text{tr}\{\mathbf{MSE}_i\}$ and I_l^{PU} can be written as $\text{tr}\{\mathbf{MSE}_i\} = \|\boldsymbol{\mu}_i\|_2^2$ and $I_l^{PU} = \|\boldsymbol{\iota}_l\|_2^2$, where $\boldsymbol{\mu}_i$ and $\boldsymbol{\iota}_l$ are given as

$$\boldsymbol{\mu}_i = \begin{bmatrix} \left(\mathbf{V}_i^T \otimes \mathbf{U}_i^H\right) vec\left(\mathbf{H}_{ii}\right) - vec\left(\mathbf{I}_{d_i}\right) \\ \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{U}_i^H\right) vec\left(\mathbf{H}_{ij}\right) \right\rfloor_{j\in\mathcal{S}, j\neq i} \\ \left\lfloor \left\lfloor \sqrt{\kappa}\left(\left(\boldsymbol{\Gamma}_\ell \mathbf{V}_j\right)^T \otimes \mathbf{U}_i^H\right) vec\left(\mathbf{H}_{ij}\right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \\ \left\lfloor \left\lfloor \sqrt{\beta}\left(\mathbf{V}_j^T \otimes \left(\mathbf{U}_i^H \boldsymbol{\Gamma}_\ell\right)\right) vec\left(\mathbf{H}_{ij}\right) \right\rfloor_{\ell\in\mathcal{D}_i^{(R)}} \right\rfloor_{j\in\mathcal{S}} \\ \sigma_i vec\left(\mathbf{U}_i\right) \end{bmatrix} \quad (54)$$

$$\boldsymbol{\iota}_l = \begin{bmatrix} \left\lfloor \left(\mathbf{V}_j^T \otimes \mathbf{I}_{T_l}\right) vec\left(\mathbf{G}_{lj}\right) \right\rfloor_{j\in\mathcal{S}} \\ \sqrt{\kappa}\left\lfloor \left\lfloor \left(\left(\boldsymbol{\Gamma}_\ell \mathbf{V}_j\right)^T \otimes \mathbf{I}_{T_l}\right) vec\left(\mathbf{G}_{lj}\right) \right\rfloor_{\ell\in\mathcal{D}_j^{(T)}} \right\rfloor_{j\in\mathcal{S}} \end{bmatrix}, \quad (55)$$

where $\mathcal{D}_j^{(R)}$ represents the set $\{1\cdots\tilde{N}_j\}$, $\mathcal{D}_j^{(T)}$ represents the set $\{1\cdots\tilde{M}_j\}$ and $\boldsymbol{\Gamma}_\ell$ is a square matrix with zero elements, except for the ℓ-th diagonal element, equal to 1. Using the vector forms (54) and (55), the problem (20) can be rewritten as

$$\begin{aligned} \min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau}} \quad & \sum\nolimits_{i\in\mathcal{S}} \tau_i && (56a) \\ \text{s.t.} \quad & \|\boldsymbol{\mu}_i\|_2^2 \leq \tau_i,\ \|\boldsymbol{\Delta}_{\boldsymbol{i}}\|_F \leq \delta_i,\ i\in\mathcal{S}, && (56b) \\ & \|vec\left(\mathbf{V}_i\right)\|_2^2 \leq P_i,\ i\in\mathcal{S}^{UL}, && (56c) \\ & \left\|\left\lfloor vec\left(\mathbf{V}_i\right)\right\rfloor_{i\in\mathcal{S}^{DL}}\right\|_2^2 \leq P_0, && (56d) \\ & \|\boldsymbol{\iota}_l\|_2^2 \leq \lambda_l,\ \|\boldsymbol{\Lambda}_{\boldsymbol{l}}\|_F \leq \theta_l,\ l=1,\ldots,L. && (56e) \end{aligned}$$

Semi-infinite optimization problems can be formulated in terms of LMIs. Such a reduction, if possible, has important practical consequences: It means that those semi-infinite problems can be solved efficiently with interior-point methods for LMI problems [42]. Note that the constraints (56b) and (56e) are not in the form of an LMI because the optimization variables do not appear linearly in these constraints. To recast the semi-infinite problem (56) as a SDP problem, the Schur complement lemma [42] is used to rewrite the constraints (56b) and (56e) in LMI form[10].

[10]Schur Complement Lemma is stated as follows: Let $\mathbf{Q}$ and $\mathbf{R}$ be symmetric matrices. Then the following two expressions are equivalent.

$$\begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^* & \mathbf{R} \end{bmatrix} \succeq \mathbf{0} \quad \triangleq \quad \mathbf{R}\succeq\mathbf{0},\ \mathbf{Q}-\mathbf{S}\mathbf{R}^{-1}\mathbf{S}^*\succeq\mathbf{0}.$$

The resulting optimization problem is written as

$$\begin{aligned} \min_{\mathbf{V},\mathbf{U},\boldsymbol{\tau}} \quad & \sum_{i\in\mathcal{S}} \tau_i && (57a) \\ \text{s.t.} \quad & \begin{bmatrix} \tau_i & \boldsymbol{\mu}_i^H \\ \boldsymbol{\mu}_i & \mathbf{I}_{A_i} \end{bmatrix} \succeq 0,\ \|\boldsymbol{\Delta}_{\boldsymbol{i}}\|_F \leq \delta_i,\ i\in\mathcal{S}, && (57b) \\ & \|vec\left(\mathbf{V}_i\right)\|_2^2 \leq P_i,\ i\in\mathcal{S}^{UL}, && (57c) \\ & \left\|\left\lfloor vec\left(\mathbf{V}_i\right)\right\rfloor_{i\in\mathcal{S}^{DL}}\right\|_2^2 \leq P_0, && (57d) \\ & \begin{bmatrix} \lambda_l & \boldsymbol{\iota}_l^H \\ \boldsymbol{\iota}_l & \mathbf{I}_{B_l} \end{bmatrix} \succeq 0,\ \|\boldsymbol{\Lambda}_{\boldsymbol{l}}\|_F \leq \theta_l,\ l=1,\ldots,L, && (57e) \end{aligned}$$

where the dimensions of the identity matrices in (57b) and (57e) are given in (22) and (23), respectively.

To further simplify the problem (57), the following lemma from [32] is used to relax the semi-infiniteness of the constraints (57b) and (57e), see also [45] for a similar approach.

Lemma 1: Given matrices $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{A}$ with $\mathbf{A} = \mathbf{A}^H$, the semi-infinite LMI of the form of

$$\mathbf{A} \succeq \mathbf{P}^H\mathbf{X}\mathbf{Q} + \mathbf{Q}^H\mathbf{X}^H\mathbf{P}, \qquad \forall\mathbf{X} : \|\mathbf{X}\|_F \leq \rho,$$

holds if and only if $\exists\epsilon \geq 0$ such that

$$\begin{bmatrix} \mathbf{A} - \epsilon\mathbf{Q}^H\mathbf{Q} & -\rho\mathbf{P}^H \\ -\rho\mathbf{P} & \epsilon\mathbf{I} \end{bmatrix} \succeq 0. \quad (58)$$

To apply Lemma 1, we need to separate the estimated channel and the channel estimation error. To this end, the LMI in (57b) is first expressed as

$$\begin{bmatrix} \tau_i & \tilde{\boldsymbol{\mu}}_i^H \\ \tilde{\boldsymbol{\mu}}_i & \mathbf{I}_{A_i} \end{bmatrix} + \begin{bmatrix} 0 & \boldsymbol{\mu}_{\Delta_i}^H \\ \boldsymbol{\mu}_{\Delta_i} & \mathbf{0}_{A_i\times A_i} \end{bmatrix} \succeq 0, \quad (59)$$

where $\tilde{\boldsymbol{\mu}}_i$ is defined in (24) and $\boldsymbol{\mu}_{\Delta_i} = \mathbf{D}_{\Delta_i} vec\left(\boldsymbol{\Delta}_{\boldsymbol{i}}\right)$, where $\mathbf{D}_{\Delta_i}$ is defined in (25). By choosing

$$\mathbf{A} = \begin{bmatrix} \tau_i & \tilde{\boldsymbol{\mu}}_i^H \\ \tilde{\boldsymbol{\mu}}_i & \mathbf{I}_{A_i} \end{bmatrix},\ \mathbf{P} = \left[\mathbf{0}_{\tilde{N}_i\tilde{M}\times 1},\ \mathbf{D}_{\Delta_i}^H\right], \quad (60)$$

$$\mathbf{X} = vec\left(\boldsymbol{\Delta}_{\boldsymbol{i}}\right),\ \mathbf{Q} = \left[-1, \mathbf{0}_{1\times A_i}\right], \quad (61)$$

and applying Lemma 1, the LMI in (57b) is relaxed as

$$\begin{bmatrix} \tau_i - \epsilon_i & \tilde{\boldsymbol{\mu}}_i^H & \mathbf{0}_{1\times\tilde{N}_i\tilde{M}} \\ \tilde{\boldsymbol{\mu}}_i & \mathbf{I}_{A_i} & -\delta_i\mathbf{D}_{\Delta_i} \\ \mathbf{0}_{\tilde{N}_i\tilde{M}\times 1} & -\delta_i\mathbf{D}_{\Delta_i}^H & \epsilon_i\mathbf{I}_{\tilde{N}_i\tilde{M}} \end{bmatrix} \succeq 0,\ i\in\mathcal{S}, \quad (62)$$

$$\epsilon_i \geq 0,\ i\in\mathcal{S}. \quad (63)$$

Using a similar procedure, the LMI in (57e) is expressed as

$$\begin{bmatrix} \lambda_l & \tilde{\boldsymbol{\iota}}_l^H \\ \tilde{\boldsymbol{\iota}}_l & \mathbf{I}_{B_l} \end{bmatrix} + \begin{bmatrix} 0 & \boldsymbol{\iota}_{\Lambda_l}^H \\ \boldsymbol{\iota}_{\Lambda_l} & \mathbf{0}_{B_l\times B_l} \end{bmatrix} \succeq 0, \quad (64)$$

where $\tilde{\boldsymbol{\iota}}_l$ is given in (26) and $\boldsymbol{\iota}_{\Lambda_l} = \mathbf{E}_{\Lambda_l} vec\left(\boldsymbol{\Lambda}_{\boldsymbol{l}}\right)$, where $\mathbf{E}_{\Lambda_l}$ is defined in (27). Then the LMI in (57e) is relaxed as

$$\begin{bmatrix} \lambda_l - \eta_l & \tilde{\boldsymbol{\iota}}_l^H & \mathbf{0}_{1\times T_l\tilde{M}} \\ \tilde{\boldsymbol{\iota}}_l & \mathbf{I}_{B_l} & -\theta_l\mathbf{E}_{\Lambda_l} \\ \mathbf{0}_{T_l\tilde{M}\times 1} & -\theta_l\mathbf{E}_{\Lambda_l}^H & \eta_l\mathbf{I}_{T_l\tilde{M}} \end{bmatrix} \succeq 0,\ l=1,\ldots,L, \quad (65)$$

$$\eta_l \geq 0,\ l=1,\ldots,L. \quad (66)$$

Using the relaxed LMIs in (62) and (65), the SDP problem, which is equivalent to (19) can be formulated as (21).

Appendix B: MSE Computation

Using (13), $\text{tr}\{\mathbf{MSE}_i\}$ can be written as

$$\begin{aligned} \text{tr}\{\mathbf{MSE}_i\} &= \text{tr}\Big\{\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)^H \\ &\quad + \mathbf{U}_i^H\boldsymbol{\Sigma}_i\mathbf{U}_i\Big\} \\ &= \text{tr}\Big\{\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i - \mathbf{I}_{d_i}\right)^H\Big\} \quad (67) \end{aligned}$$

$$+\sum\nolimits_{j\in\mathcal{S},j\neq i}\mathrm{tr}\left\{\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ij}^H\mathbf{U}_i\right\}+\sigma_i^2\mathrm{tr}\left\{\mathbf{U}_i^H\mathbf{U}_i\right\}$$
$$+\sum\nolimits_{j\in\mathcal{S}}\sum\nolimits_{\ell\in\mathcal{D}_j^{(T)}}\kappa\mathrm{tr}\left\{\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{\Gamma}_\ell\mathbf{V}_j\mathbf{V}_j^H\mathbf{\Gamma}_\ell^H\mathbf{H}_{ij}^H\mathbf{U}_i\right\}$$
$$+\sum\nolimits_{j\in\mathcal{S}}\sum\nolimits_{\ell\in\mathcal{D}_i^{(R)}}\beta\mathrm{tr}\left\{\mathbf{U}_i^H\mathbf{\Gamma}_\ell\mathbf{H}_{ij}\mathbf{V}_j\mathbf{V}_j^H\mathbf{H}_{ij}^H\mathbf{\Gamma}_\ell^H\mathbf{U}_i\right\},$$

where $\mathcal{D}_j^{(R)}$ represents the set $\{1\cdots\tilde{N}_j\}$, $\mathcal{D}_j^{(T)}$ represents the set $\{1\cdots\tilde{M}_j\}$ and $\mathbf{\Gamma}_\ell$ is a square matrix with zero elements, except for the ℓ-th diagonal element, equal to 1. Applying the $vec(\cdot)$ operation, and the identity $\|vec\left(\mathbf{A}\right)\|_2^2=\mathrm{tr}\left\{\mathbf{A}\mathbf{A}^H\right\}$, (67) can be rewritten as

$$\mathrm{tr}\{\mathbf{MSE}_i\}=\left\|vec\left(\mathbf{U}_i^H\mathbf{H}_{ii}\mathbf{V}_i\right)-vec(\mathbf{I}_{d_i})\right\|_2^2 \quad (68)$$
$$+\sum\nolimits_{j\in\mathcal{S},j\neq i}\left\|vec\left(\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{V}_j\right)\right\|_2^2+\sigma_i^2\left\|vec\left(\mathbf{U}_i\right)\right\|_2^2$$
$$+\sum\nolimits_{j\in\mathcal{S}}\sum\nolimits_{\ell\in\mathcal{D}_j^{(T)}}\kappa\left\|vec\left(\mathbf{U}_i^H\mathbf{H}_{ij}\mathbf{\Gamma}_\ell\mathbf{V}_j\right)\right\|_2^2$$
$$+\sum\nolimits_{j\in\mathcal{S}}\sum\nolimits_{\ell\in\mathcal{D}_i^{(R)}}\beta\left\|vec\left(\mathbf{U}_i^H\mathbf{\Gamma}_\ell\mathbf{H}_{ij}\mathbf{V}_j\right)\right\|_2^2.$$

Using the identity $vec(\mathbf{ABC})=\left(\mathbf{C}^T\otimes\mathbf{A}\right)vec\left(\mathbf{B}\right)$, (68) can be written as $\|\boldsymbol{\mu}_i\|_2^2$, where $\boldsymbol{\mu}_i$ is given in (54).

Similar to (68), I_l^{PU} can be written as

$$I_l^{PU}=\sum\nolimits_{j\in\mathcal{S}}\left(\left\|vec\left(\mathbf{G}_{lj}\mathbf{V}_j\right)\right\|_2^2+\sum\nolimits_{\ell\in\mathcal{D}_j^{(T)}}\kappa\left\|vec\left(\mathbf{G}_{lj}\mathbf{\Gamma}_\ell\mathbf{V}_j\right)\right\|_2^2\right). \quad (69)$$

Using the identity $vec(\mathbf{ABC})=\left(\mathbf{C}^T\otimes\mathbf{A}\right)vec\left(\mathbf{B}\right)$, (69) can be written as $\|\boldsymbol{\iota}_l\|_2^2$, where $\boldsymbol{\iota}_l$ is given in (55).

REFERENCES

[1] A. C. Cirik, S. Biswas, O. Taghizadeh, A. Liu, T. Ratnarajah, "Robust transceiver design in full-duplex MIMO cognitive radios," *IEEE Int. Conf. Commun. (ICC)*, pp. 1-7, May 2016.

[2] Chih-Lin I, C. Rowell, S. Han, Z. Xu, G. Li and Z. Pan, "Towards green and soft: A 5G perspective," *IEEE Comm. Mag.*, Feb. 2014.

[3] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[4] D. Bharadia and S. Katti, "Full duplex MIMO radios," *USENIX NSDI*, pp. 359-372, 2014.

[5] Y. Hua, Y. Ma, A. Gholian, Y. Li, A. C. Cirik, P. Liang, "Radio self-interference cancellation by transmit beamforming, all-analog cancellation and blind digital tuning," *Elsevier Signal Process.*, vol. 108, pp. 322-340, Mar. 2015.

[6] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," *IEEE Trans. Vel. Technol.*, vol. 62, no. 9, pp. 4494-4510, Nov. 2013.

[7] 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, July 2012.

[8] W. Li, J. Lilleberg, and K. Rikkinen, "On rate region analysis of half- and full-duplex OFDM communication links," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 9, pp. 1688-1698, Sept. 2014.

[9] A. C. Cirik, S. Biswas, S. Vuppala and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channelsQoS and energy-efficiency considerations," *IEEE Trans. Communications*, vol. 64, no. 11, pp. 4635-4651, Nov. 2016.

[10] A. C. Cirik, O. Taghizadeh, L. Lampe, R. Mathar and Y. Hua, "Linear transceiver design for full-duplex multi-cell MIMO systems," *IEEE Access*, vol. 4, pp. 4678-4689, 2016.

[11] A. C. Cirik, R. Wang, Y. Rong and Y. Hua, "MSE based transceiver designs for bi-directional full-duplex MIMO systems," in *Proc. IEEE SPAWC*. pp. 384-388, 2014.

[12] A. C. Cirik, R. Wang, Y. Rong, and Y. Hua,"MSE-based transceiver designs for full-duplex MIMO cognitive radios," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2056-2070, Jun. 2015.

[13] T. M. Kim, H. J. Yang, and A. Paulraj, "Distributed sum-rate optimization for full-duplex MIMO system under limited dynamic range," *IEEE Signal Process. Letters*, vol. 20, no. 6, pp. 555-558, Jun. 2013.

[14] D. Nguyen, L. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896-4910, Sept. 2014.

[15] S. Li, R. Murch, and V. Lau, "Linear transceiver design for full-duplex multi-user MIMO system," in *Proc. IEEE ICC*, pp. 4921-4926, Jun. 2014.

[16] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201-220, Feb. 2005.

[17] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79-89, May 2007.

[18] G. Zheng, S. Ma, K. Kit Wong, and T.-S. Ng, "Robust beamforming in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 9, no. 2, pp. 570-576, Feb. 2010.

[19] L. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor, "Robust cognitive beamforming with partial channel state information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 8, pp. 4143-4153, Aug. 2009.

[20] Y. Zhang, E. DallAnese, and G. B. Giannakis, "Distributed optimal beamformers for cognitive radios robust to channel uncertainties," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6495-6508, Dec. 2012.

[21] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1183-1201, Mar. 2011.

[22] G. Zheng, K. K. Wong, and B. Ottersten, "Robust cognitive beamforming with bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4871-4881, Dec. 2009.

[23] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust linear transceiver design in MIMO ad hoc cognitive radio networks with imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1448-1457, May 2011.

[24] D. P. Bertsekas, "Nonlinear Programming", 2nd ed. Belmont, MA: Athena Scientific, 1999.

[25] A. Mutapcic and S. Boyd, "Cutting-set methods for robust convex optimization with pessimizing oracles," *Optim. Methods Softw.*, vol. 24, no. 3, pp. 381-406, Jun. 2009.

[26] A. Mutapcic, S. Kim, and S. Boyd, "A tractable method for robust downlink beamforming in wireless communications," in *Proc. ACSSC*, pp. 1224-1228, Nov. 2007.

[27] S. Huberman and T. Le-Ngoc, "Full-duplex MIMO precoding for sum-rate maximization with sequential convex programming," *IEEE Trans. Wireless Commun.*, vol. 64, no. 11, pp. 5103-5112, Nov. 2015.

[28] 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, Apr. 2015.

[29] P. Wu, R. Schober and V. K. Bhargava, "Robust transceiver design for SC-FDE multi-hop full-duplex decode-and-forward relaying systems," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1129-1145, Feb. 2016.

[30] J. Zhang, O. Taghizadeh, and M. Haardt, "Robust transmit beamforming design for full-duplex point-to-point MIMO systems," in *Proc. ISWCS*, pp. 1-5, 2013.

[31] R. Feng, Q. Li, Q. Zhang, and J. Qin, "Robust secure beamforming in MISO full-duplex two-way secure communications," *IEEE Trans. Veh. Technol.*, vol. 65, no. 1, pp. 408-414, Jan. 2016.

[32] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared-error estimation in the presence of model uncertainties," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 161-176, Jan. 2005.

[33] 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, Aug. 2008.

[34] 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, Jan. 2005.

[35] G. Scutari, D. P. Palomar, and S. Barbarossa, "Cognitive MIMO radio," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 46-59, Nov. 2008.

[36] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Flexible design of cognitive radio wireless systems: From game theory to variational inequality theory," *IEEE Signal Proc. Mag.*, vol. 26, no. 5, pp. 107-123, Sept. 2009.

[37] G. Scutari and D. P. Palomar, "MIMO cognitive radio: A game-theoretical approach," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 761-780, Feb. 2010.

[38] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117-3131, May 2011.

[39] O. Simeone, U. Spagnolini and Y. Bar-Ness, "Linear and non-linear precoding/decoding for MIMO systems using the fading correlation at the transmitter," in *Proc. IEEE SPAWC*, pp. 6-10, Jun. 2003.

[40] N. I. Miridakis and D. D. Vergados, "A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems," *IEEE Commun. Surveys and Tut.*, vol. 15, no. 1, pp. 312-335, 2013.

[41] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.

[42] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[43] M. Botros Shenouda and T. N. Davidson, "On the design of linear transceivers for multiuser systems with channel uncertainty," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 1015-1024, Aug. 2008.

[44] E. A. Gharavol and E. G. Larsson, "The sign-definiteness lemma and its applications to robust transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 238-252, Jan. 2013.

[45] J. Jose, N. Prasad, M. Khojastepour and S. Rangarajan, "On robust weighted-sum rate maximization in MIMO interference networks," in *Proc. IEEE ICC*, pp. 1-6, Jun. 2011.

[46] K. Phan, S. Vorobyov, N. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-Aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2323-2335, Jun. 2009.

[47] T. W. Ban, W. Choi, B. C. Jung, and D. K. Sung, "Multi-user diversity in a spectrum sharing system," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 102-106, Jan. 2009.

[48] A. Ghasemi and E. S. Sousa, "Fundamental limits of spectrum-sharing in fading environments," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 649-658, Feb. 2007.

[49] D. I. Kim, L. B. Le, and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5517-5527, Dec. 2008.

[50] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic spectrum access via periodic channel sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 785-796, Feb. 2008.

[51] F. Gao, R. Zhang, Y.-C. Liang, and X. Wang, "Multi-antenna cognitive radio systems: Environmental learning and channel training," in *Proc. IEEE ICASSP*, Apr. 2009, pp. 2329-2332.

[52] E. A. Gharavol, Y.-C. Liang, and K. Mouthaan, "Robust downlink beamforming in multiuser MISO cognitive radio networks with imperfect channel-state information," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2852-2860, Jul. 2010.

[53] J. M. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10-12, Feb. 2005.

[54] P. Ubaidulla and A. Chockalingam, "Relay precoder optimization in MIMO-relay networks with imperfect CSI," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5473-5484, Nov. 2011.

[55] J. Liu, F. Gao, and Z. Qiu, "Robust transceiver design for downlink multiuser MIMO AF relay systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2218-2231, Apr. 2015.

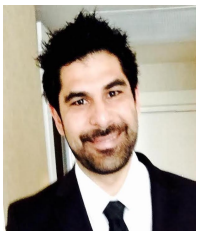
[56] P. Ubaidulla and S. Aissa, "Robust two-way cognitive relaying: Precoder designs under interference constraints and imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 13, no. 5, pp. 2478-2489, May 2014.

[57] M.-M. Zhao, Y. Cai, B. Champagne, and M. Zhao, "Min-Max MSE transceiver with switched preprocessing for MIMO interference channels," in *Proc. IEEE PIMRC*, pp. 198-202, Sep. 2014.

[58] H. Shen, B. Li, M. Tao, and X. Wang, "MSE-based transceiver designs for the MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480-3489, Nov. 2010.

[59] 3GPP, TR 36.828, "Further enhancements to LTE time division duplex (TDD) for downlink-uplink (DL-UL) interference management and traffic adaptation (Release 11)," Jun. 2012.

[60] S. Goyal, P. Liu, S. Panwar, R. DiFazio, R. Yang, J. Li, and E. Bala, "Improving small cell capacity with common-carrier full duplex radios," in *Proc. IEEE ICC 2014*, pp. 4987-4993, Jun. 2014.



Ali Cagatay Cirik (S'13-M'14) received the B.S and M.S. degrees in telecommunications and electronics engineering from Sabanci University, Istanbul, Turkey, in 2007 and 2009, respectively, and Ph.D. degree in electrical engineering from University of California, Riverside in 2014. He held research fellow positions at Centre for Wireless Communications, Oulu, Finland and Institute for Digital Communications (IDCOM), University of Edinburgh, U.K between June 2014 and November 2015. His industry experience includes internships at Mitsubishi Electric Research Labs (MERL), Cambridge, MA, in 2012 and at Broadcom Corporation, Irvine, CA, in 2013. Currently, he is working as a research scientist at Sierra Wireless, Richmond, Canada. He is also affiliated with University of British Columbia, Vancouver, Canada. His primary research interests are full-duplex communication, non-orthogonal multiple-access (NOMA), MIMO signal processing, and convex optimization.

Sudip Biswas (S'16-M'17) received the B.Tech. degree in electronics and communication engineering from the Sikkim Manipal Institute of Technology, Sikkim, India, in 2010, the M.Sc. degree in signal processing and communications from the University of Edinburgh, Edinburgh, U.K., in 2013 and the Ph.D. degree in digital communications at the University of Edinburgh's Institute for Digital Communications in 2017. Currently he is working as a research scientist at the Institute for Digital Communications (IDCOM), University of Edinburgh. His research interests include various topics in wireless communications and network information theory with particular focus on stochastic geometry and possible 5G technologies such as massive MIMO, mmWave, full-duplex and NOMA.

Omid Taghizadeh received his M.Sc. degree in Communications and Signal Processing in April 2013, from Ilmenau University of Technology, Ilmenau, Germany. From September 2013, he has been a research assistant in the Institute for Theoretical Information Technology, RWTH Aachen University. His research interests include full-duplex wireless systems, MIMO communications, optimization, and resource allocation in wireless networks.

Tharmalingam Ratnarajah (A'96-M'05-SM'05) is currently with the Institute for Digital Communications, University of Edinburgh, Edinburgh, UK, as a Professor in Digital Communications and Signal Processing and the Head of Institute for Digital Communications. His research interests include signal processing and information theoretic aspects of 5G and beyond wireless networks, full-duplex radio, mmWave communications, random matrices theory, interference alignment, statistical and array signal processing and quantum information theory. He has published over 330 publications in these areas and holds four U.S. patents. He was the coordinator of the FP7 projects ADEL (3.7M€) in the area of licensed shared access for 5G wireless networks and HARP (3.2M€) in the area of highly distributed MIMO and FP7 Future and Emerging Technologies projects HIATUS (2.7M€) in the area of interference alignment and CROWN (2.3M€) in the area of cognitive radio networks. Dr Ratnarajah is a Fellow of Higher Education Academy (FHEA), U.K..