



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Transceiver Design in Full-Duplex MIMO Cognitive Radios under Channel Uncertainties

Citation for published version:

Cagatay Cirik, A, Filippou, MC & Ratnarajah, T 2016, 'Transceiver Design in Full-Duplex MIMO Cognitive Radios under Channel Uncertainties', *IEEE Transactions on Cognitive Communications and Networking*, vol. 2, no. 1, pp. 2332-7731. <https://doi.org/10.1109/TCCN.2016.2572691>

Digital Object Identifier (DOI):

[10.1109/TCCN.2016.2572691](https://doi.org/10.1109/TCCN.2016.2572691)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

IEEE Transactions on Cognitive Communications and Networking

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Transceiver Design in Full-Duplex MIMO Cognitive Radios under Channel Uncertainties

Ali Cagatay Cirik, *Member, IEEE*, Miltiades C. Filippou, *Member, IEEE*, and Tharmalingam Ratnarajah, *Senior Member, IEEE*

Abstract—In this work, the operation of a full-duplex (FD) multiple-input multiple-output (MIMO) underlay cognitive radio (CR) network is studied. Each pair of secondary users (SUs) operates in FD mode and each SU communicates with its peer, within the coverage area of multiple primary users (PUs). Each SU suffers from self-interference, along with interference created by all other SUs and, in its turn, generates interference towards the PU. We assume that the channel-state-information (CSI) of the interference links between the SUs and the PU is imperfectly known and the channel estimation errors are norm bounded. Under such channel uncertainties, we address the problem of robust minimization of the aggregate mean-squared-error (MSE) of all estimated symbols, subject to a set of transmit power constraints, as well as interference constraints, with the aim of protecting PU communication. It is shown that the problem can be cast as a semidefinite programming (SDP) problem and the optimal precoding matrices can be obtained via an iterative algorithm. By means of simulation, the proposed FD precoding scheme is shown to outperform the standard half-duplex (HD) scheme.

Keywords—Cognitive radio, full-duplex, imperfect CSI, MIMO, transceiver design.

I. INTRODUCTION

COGNITIVE radio (CR) is a promising technology to enhance the spectrum efficiency utilization by allowing unlicensed secondary-users (SUs) operate within the service range of licensed primary-users (PUs), while causing tolerable interference to the PUs. Traditionally, the secondary network is deployed as half-duplex (HD), in which transmission and reception are orthogonal in time or frequency [1]-[14]. Among the emerging technologies for next-generation wireless networks, full-duplex (FD) communication is envisioned as a way of potentially doubling the throughput of wireless communication systems, since it enables available spectral resources to be fully utilized both in time and frequency [15],

[16]. However, most related work has dealt with modeling and cancellation of the high-power self-interference signal which is the leakage of transmit signal to the collocated receiver, and is the fundamental challenge in implementing a FD radio. Fortunately, many feasible solutions including antenna, analog and digital cancellation methods have been demonstrated experimentally to mitigate the overwhelming self-interference [17]-[19], and have made FD communication more practical in recent years. However, various practical implementation issues, such as protocol and resource allocation algorithm design, need to be reinvestigated in the context of FD communications. Particularly, the deployment of FD systems in a network setting needs to be addressed through power and interference management policies.

Transceiver designs for the sum-rate maximization problem in FD multiple-input multiple-output (MIMO) non-cognitive-radio systems have been investigated in [20]-[27], but the authors in [20]-[25] have not considered the fundamental impediment of FD radios, i.e., the limited dynamic-range (DR) caused by non-ideal amplifiers, oscillators, analog-to-digital converters (ADCs), digital-to-analog converters (DACs), etc [28] in their derivations. Although FD communication has drawn significant research interest, research on FD CR systems is still in its infancy, and has not been thoroughly studied in the literature so far. Research problems related to interweave FD CR systems have been investigated in [29]-[31]. In particular, the authors have explored FD techniques at the SUs to achieve simultaneous spectrum sensing and data transmission to significantly improve sensing performance while increasing data transmission efficiency. Another emerging research trend concerns underlay cooperative relaying systems [32]-[34] where FD cognitive radios are employed. A sum mean-squared-error (MSE) minimization problem for a FD MIMO cellular and ad-hoc underlay CR system has been studied in [35], in which the optimization problem has been cast as a second-order-cone-program (SOCP). The authors assume that perfect channel-state-information (CSI) is available at the transmitters, which is practically impossible due to the inaccurate channel estimation and the lack of full SU-PU cooperation. Therefore, robust transceiver designs that take into account imperfect channel knowledge are of interest, which have not been reported (to the best of our knowledge) so far for FD underlay CR systems. Note that the SOCP-based algorithm proposed in [35] for the FD CR system cannot be applied under CSI with norm-bounded uncertainties.

Motivated by the above, in this paper we propose an iterative transceiver design scheme for a MIMO FD system, by focusing

This work is supported by the Seventh Framework Programme for Research of the European Commission under grant number ADEL-619647.

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

T. Ratnarajah is with the Institute for Digital Communications, School of Engineering, University of Edinburgh, Edinburgh EH9 3FG, U.K. (email: t.ratnarajah@ed.ac.uk).

M. C. Filippou was with the Institute for Digital Communications, University of Edinburgh, Edinburgh EH9 3FG, U.K. He is now with Intel Germany GmbH, Am Campeon 10-12, 85579, Neubiberg, Germany (e-mail: miltiades.filippou.de@ieee.org)

on an underlay (spectrum sharing) CR interference channel (i.e., an ad-hoc FD CR network). In such system scenarios, the SUs have to use their resources (i.e., power, antennas), in a way that the received interference power by the PU is below a predefined Quality-of-Service (QoS) threshold. We take into account the limited-DR of the transmitters and receivers caused by the FD operation. Since the availability of perfect CSI, regarding the interference links between the SUs and the PUs is deemed highly unrealistic, we also consider a norm-bounded-error model, i.e., the instantaneous channel lies in a known set of possible values, which represents the amount of uncertainty on the channel [1]-[3].

According to the investigated system model, we consider K pairs of FD SUs which exchange information, while, at the same time, protect primary communication. Each SU suffers not only from self-interference due to its operation in the FD mode, but also from co-channel-interference (CCI) due to concurrent transmissions from all SUs. The channels between SUs are assumed to be perfectly known, however, the channels between SUs and PUs are subject to estimation errors. More particularly, our contributions are as follows:

- Prior work on FD systems assumes that perfect CSI is available at the transmitters, which is practically impossible due to the inaccurate channel estimation. Therefore, robust transceiver designs that take into account imperfect channel knowledge are of interest, which have not been reported (to the best of our knowledge) so far for full-duplex cellular systems or interference channels. This work tries to fill this gap and reveals useful insights into FD CR systems via MSE-based optimization under imperfect CSI, and aims at showing the importance of accurate channel estimation, and how critical it is for successful deployment of FD systems.
- We consider the sum-MSE of the estimated symbols as the objective function to be minimized, subject to a set of transmit power constraints at the SUs and interference power constraints at the PU from each SU. Given that the problem is semi-infinite and non-convex, an iterative scheme is proposed, with the aim of jointly designing the transceiver (beamforming) matrices at the transmitter/receiver of each SU. First, the semi-infinite constraints are transformed into tractable forms, and then, a local convex approximation technique is adopted. The resulting problem can be solved via an iterative semidefinite programming (SDP) algorithm, the convergence of which is guaranteed to a stationary point. Due to the transmit and receive distortion at the FD nodes, i.e., RF hardware impairments, which are the major impediment to FD systems, MSE is a complicated function, which makes the transformation of the non-convex optimization problem into an SDP problem complicated.
- Instead of considering individual interference power constraints at the PU from each SU, inspired from [3], we extend the proposed algorithm to an aggregate interference power constraint induced by all SUs to the PU, and apply a primal decomposition technique to allocate the total interference among SUs dynamically.

- Moreover, we show that the proposed CR underlay algorithm is not only applicable to FD MIMO interference channels, but also to FD cellular systems, in which a base station (BS) operating in FD mode simultaneously serves multiple HD uplink (UL) and downlink (DL) users. In this setup, in addition to self-interference channel at the BS, the difficulty of the design problem is increased further by the CCI caused by the users in the UL channel to those in the DL channel.
- Our joint transceiver design scheme is numerically evaluated and it is shown that significant throughput improvement is achieved over the corresponding HD MIMO CR system. Moreover, the importance of channel estimation in FD systems is shown.

A. Rationale for MSE-Based Optimizations

MSE-based transceiver designs have been considered extensively due to the decent performance and the significantly reduced complexity of this metric. It has been shown in [36] that minimum mean-squared-error (MMSE) estimation plays an important role in approaching the information-theoretic limits of Gaussian channels. When MMSE receiver is used, MSE-based optimization problems are equivalent to signal-to-interference-plus-noise ratio (SINR)-based optimization problems, since they are related as [37],

$$\text{MSE} = \frac{1}{1 + \text{SINR}}. \quad (1)$$

Therefore, rate-based optimization using $\log_2(1 + \text{SINR})$ can be conveniently transformed into MSE-based optimization, $-\log_2(\text{MSE})$. And as mentioned in [38], the user-wise MSE can be used to approximate the achievable rate of the users when they jointly decode their streams. With (1), instead of considering each design criterion such as the MSE and the maximal mutual information in a separate way, a unifying framework can be developed. The link between most practical objective functions and the main diagonal elements of the MSE matrix has been established in [37] for point-to-point multicarrier MIMO communications, and this work has been extended to multicarrier MIMO relay communications in [39]. Our literature survey reveals that MSE-based optimization problems have been considered for many communication systems but not for robust FD CR systems. This work tries to fill this gap and reveals useful insights into FD CR systems via MSE-based optimization under imperfect CSI.

Notation: The following notations are used in this paper. Matrices and vectors are denoted as bold capital and lowercase letters, respectively. $(\cdot)^H$ is the conjugate transpose operator. $\mathbb{E}\{\cdot\}$ stands for the statistical expectation of a random variable; \mathbf{I}_N is the N by N identity matrix; $\mathbf{0}_{N \times M}$ is the N by M zero matrix; $\text{tr}\{\cdot\}$ is the trace; $\text{diag}(\mathbf{A})$ is the diagonal matrix with the same diagonal elements as \mathbf{A} . $\mathcal{CN}(\mu, \sigma^2)$ denotes a complex Gaussian distribution with mean μ and variance σ^2 . $\text{vec}(\cdot)$ stacks the elements of a matrix to one long column vector. The operator \otimes denotes Kronecker product and \perp denotes the statistical independence. $\|\mathbf{X}\|_F$ and $\|\mathbf{x}\|_2$ denote the Frobenius norm of a matrix \mathbf{X} and the Euclidean norm of

a vector \mathbf{x} , respectively. Finally, $\Re\{\mathbf{X}\}$ denotes the real part of \mathbf{X} . $\mathbf{A} \succeq \mathbf{0}$ indicates that \mathbf{A} is a positive semidefinite matrix.

II. SYSTEM MODEL

In this section, we describe the system model of a FD MIMO CR system, in which K pairs of FD SUs communicate simultaneously within the service range of L PUs as seen in Fig. 1. Let us denote the set of SU pairs and PUs with $\mathcal{K} \triangleq \{1, \dots, K\}$ and $\mathcal{L} \triangleq \{1, \dots, L\}$, respectively. We assume that each SU node that belongs to the i th pair, is equipped with N_i and M_i transmit and receive antennas, respectively.

A. Signal and channel model

In what follows, $i^{(a)}$ denotes SU $a \in \{1, 2\}$ belonging to pair $i \in \mathcal{K}$. The SU $i^{(a)}$, $i \in \mathcal{K}$, $a \in \{1, 2\}$ receives signals from all the SU transmitters in the system via MIMO channels. $\mathbf{H}_{ii}^{(ab)} \in \mathbb{C}^{M_i \times N_i}$ is the desired channel between the transmitter of node b , where $b \in \{1, 2\}$, $b \neq a$ and the receiver of node a , when both nodes (SUs) belong to the i -th pair. $\mathbf{H}_{ii}^{(aa)} \in \mathbb{C}^{M_i \times N_i}$, $a \in \{1, 2\}$ denotes the self-interference channel of the SU $i^{(a)}$. Also, $\mathbf{H}_{ij}^{(ac)} \in \mathbb{C}^{M_i \times N_j}$, $(a, c) \in \{1, 2\}$ denotes the CCI channel from the transmit antennas of the SU c in the j th SU pair to the receive antennas of SU a in the i th pair, $(i, j) \in \mathcal{K}$ and $j \neq i$. All the channel matrices are assumed to be mutually independent, and the entries of each matrix are independent and identically distributed (i.i.d.) circular complex Gaussian variables with zero mean, independent real and imaginary parts, each with variance $1/2$.

The transmitted data streams of size d_i at the SU $i^{(a)}$ are denoted as $\mathbf{d}_i^{(a)} \in \mathbb{C}^{d_i}$, $i \in \mathcal{K}$, $a \in \{1, 2\}$, and are assumed to be complex, zero mean, i.i.d. with unit variance. The $N_i \times 1$ signal vector transmitted by the SU $i^{(a)}$ is given by

$$\mathbf{x}_i^{(a)} = \mathbf{V}_i^{(a)} \mathbf{d}_i^{(a)}, \quad i \in \mathcal{K}, a \in \{1, 2\}, \quad (2)$$

where $\mathbf{V}_i^{(a)} \in \mathbb{C}^{N_i \times d_i}$ represents the transmit beamforming matrix applied at the node $i^{(a)}$.

According to the investigated system model, we consider a FD MIMO interference channel between SUs that suffers from self-interference and CCI from other pairs. Thus, the SU $i^{(a)}$ receives a combination of the signals transmitted by all the transmitters along with additive noise. The $M_i \times 1$ received signal at the SU $i^{(a)}$ is written as

$$\begin{aligned} \mathbf{y}_i^{(a)} &= \sqrt{\rho_i} \mathbf{H}_{ii}^{(ab)} \left(\mathbf{x}_i^{(b)} + \mathbf{c}_i^{(b)} \right) + \underbrace{\sqrt{\eta_{ii}} \mathbf{H}_{ii}^{(aa)} \left(\mathbf{x}_i^{(a)} + \mathbf{c}_i^{(a)} \right)}_{\text{Self-Interference}} \\ &+ \underbrace{\sum_{j \neq i} \sum_{c=1}^2 \sqrt{\eta_{ij}^{(ac)}} \mathbf{H}_{ij}^{(ac)} \left(\mathbf{x}_j^{(c)} + \mathbf{c}_j^{(c)} \right)}_{\text{CCI}} + \mathbf{e}_i^{(a)} \\ &+ \mathbf{n}_i^{(a)}, \quad i \in \mathcal{K}, (a, b) \in \{1, 2\} \text{ and } a \neq b. \end{aligned} \quad (3)$$

Here, $\mathbf{n}_i^{(a)} \in \mathbb{C}^{M_i}$ is the additive white gaussian noise (AWGN) vector at SU $i^{(a)}$ with zero mean and covariance matrix \mathbf{I}_{M_i} , and it is uncorrelated to all the transmitted

signals¹. In (3), ρ_i denotes the average power gain of the i th SU transmitter-receiver pair, η_{ii} denotes the average power gain of the self-interference channel at the i th SU pair, and $\eta_{ij}^{(ac)}$ denotes the average power gain of the CCI channel between the nodes at the $i^{(a)}$ th and $j^{(c)}$ th SU pair.²

Furthermore, in (3), $\mathbf{c}_i^{(a)} \in \mathbb{C}^{N_i}$, $i \in \mathcal{K}$, $a \in \{1, 2\}$ is the noise at the transmit antennas of SU $i^{(a)}$, which models the effect of limited transmitter DR, and closely approximates the effects of additive power-amplifier noise, non-linearities in the DAC and phase noise. The covariance matrix of $\mathbf{c}_i^{(a)}$ is given by κ times the energy of the intended signal at each transmit antenna [28]. In particular $\mathbf{c}_i^{(a)}$ can be modeled as

$$\mathbf{c}_i^{(a)} \sim \mathcal{CN} \left(\mathbf{0}_{N_i}, \kappa \text{diag} \left(\mathbf{V}_i^{(a)} \left(\mathbf{V}_i^{(a)} \right)^H \right) \right), \quad \mathbf{c}_i^{(a)} \perp \mathbf{x}_i^{(a)}. \quad (4)$$

Finally, in (3), $\mathbf{e}_i^{(a)} \in \mathbb{C}^{M_i}$, $i \in \mathcal{K}$, $a \in \{1, 2\}$ is the additive receiver distortion at the receive antennas of the SU $i^{(a)}$, which models the effect of limited receiver DR, and closely approximates the combined effects of additive gain-control noise, non-linearities in the ADC and phase noise. The covariance matrix of $\mathbf{e}_i^{(a)}$ is given by β times the energy of the undistorted received signal at each receive antenna [28]. In particular, $\mathbf{e}_i^{(a)}$ can be modeled as

$$\mathbf{e}_i^{(a)} \sim \mathcal{CN} \left(\mathbf{0}_{M_i}, \beta \text{diag} \left(\Phi_i^{(a)} \right) \right), \quad \mathbf{e}_i^{(a)} \perp \mathbf{u}_i^{(a)}, \quad (5)$$

where $\Phi_i^{(a)} = \text{Cov}\{\mathbf{u}_i^{(a)}\}$ and $\mathbf{u}_i^{(a)}$ is the undistorted received signal vector at SU $i^{(a)}$, i.e., $\mathbf{u}_i^{(a)} = \mathbf{y}_i^{(a)} - \mathbf{e}_i^{(a)}$.

This transmitter/receiver distortion model is valid, since it was shown by hardware measurements in [40] and [41] that the non-ideality of the transmitter and receiver chain can be approximated by an independent Gaussian noise model, respectively. Note that this model has also been commonly used in [26]-[27], [35], [42]-[44].

B. Self-interference cancellation

We assume that SU $i^{(a)}$ knows the self-interfering code-words $\mathbf{x}_i^{(a)}$, and its self-interference channel $\mathbf{H}_{ii}^{(aa)}$, so the self-interference term $\sqrt{\eta_{ii}} \mathbf{H}_{ii}^{(aa)} \mathbf{x}_i^{(a)}$ is known, and thus can

¹Since the SU receiver cannot distinguish the interference generated by the PUs from the background thermal noise, the noise vector $\mathbf{n}_i^{(a)}$ in (3) captures the background thermal noise as well as the interference generated by the PUs, possibly after prewhitening. In particular, we assume that the PUs' sum-interference is estimated and measured at the receiving node of SUs [8]. To do so, the SUs need to be "cognitive users" which are aware of the environment [9]. To achieve that, the protocol for SUs can be designed as follows: every frame contains sensing sub-frame and data transmission sub-frame. During the sensing sub-frame, all SU transmitters remain silent, and thus the SU receivers can measure the effect from the PUs and background noise [1], [9]. This assumption that noise term includes both thermal noise and PU sum-interference is also adopted in [1]-[4].

²Note that in (3), the power gains ρ and η correspond to the large-scale fading factors, which are distance-based, therefore they are assumed to be constant from time-slot to time-slot, since mobility is not taken into account in the studied scenario. On the contrary, channels \mathbf{H} are considered to model the fast fading phenomena. Also note that the variance of the noise does not influence the algorithms discussed later.

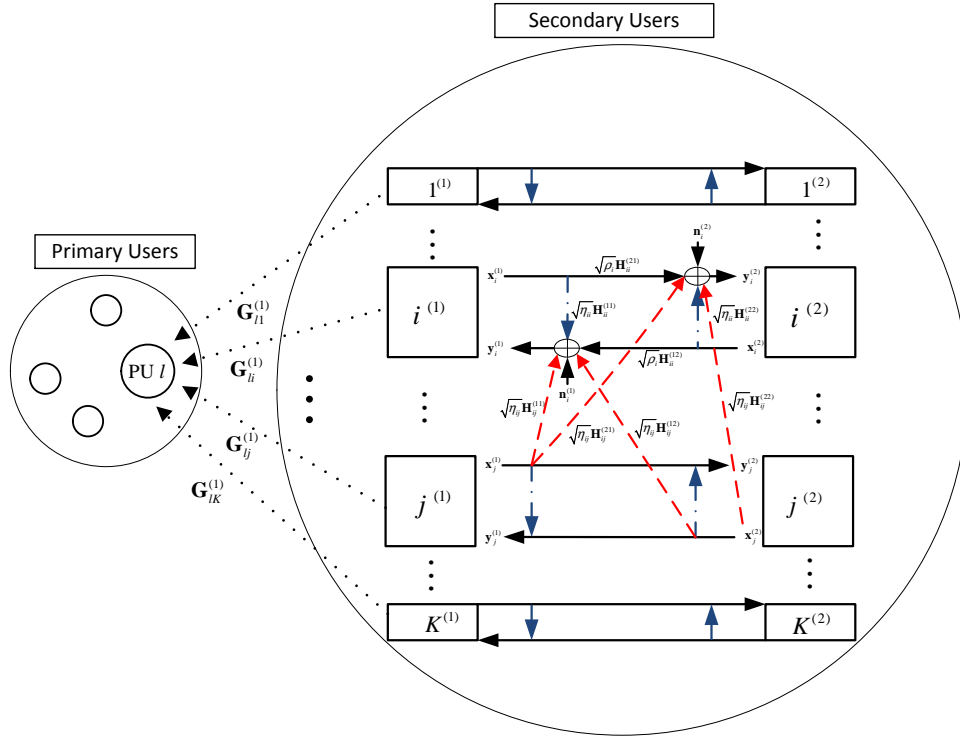


Fig. 1. Full-duplex MIMO cognitive radio system.

be canceled [28]. The received signal after self-interference cancellation can then be written as

$$\begin{aligned}\tilde{\mathbf{y}}_i^{(a)} &= \mathbf{y}_i^{(a)} - \sqrt{\eta_i} \mathbf{H}_{ii}^{(aa)} \mathbf{x}_i^{(a)} \\ &= \sqrt{\rho_i} \mathbf{H}_{ii}^{(ab)} \mathbf{x}_i^{(b)} + \tilde{\mathbf{n}}_i^{(a)},\end{aligned}\quad (6)$$

where $\tilde{\mathbf{n}}_i^{(a)} \in \mathbb{C}^{M_i \times 1}$ is the residual interference components of (6) after self-interference cancellation and is given by

$$\begin{aligned}\tilde{\mathbf{n}}_i^{(a)} &= \sqrt{\rho_i} \mathbf{H}_{ii}^{(ab)} \mathbf{c}_i^{(b)} + \sqrt{\eta_i} \mathbf{H}_{ii}^{(aa)} \mathbf{c}_i^{(a)} + \mathbf{e}_i^{(a)} + \mathbf{n}_i^{(a)} \\ &+ \sum_{j \neq i} \sum_{c=1}^2 \sqrt{\eta_{ij}^{(ac)}} \mathbf{H}_{ij}^{(ac)} (\mathbf{x}_j^{(c)} + \mathbf{c}_j^{(c)}).\end{aligned}\quad (7)$$

Using (4)-(5), similar to [28], $\Sigma_i^{(a)}$, the covariance matrix of $\tilde{\mathbf{n}}_i^{(a)}$, can be approximated under $\kappa \ll 1$ and $\beta \ll 1$, i.e., by ignoring the terms including the multiplication $\kappa\beta$, as in (8) shown at the bottom of the following page.

We assume that the SU $i^{(a)}$ applies the linear receiver $\mathbf{R}_i^{(a)} \in \mathbb{C}^{d_i \times M_i}$ to estimate the signal transmitted from SU $i^{(b)}$, i.e., $\mathbf{d}_i^{(b)}$. That is

$$\begin{aligned}\hat{\mathbf{d}}_i^{(b)} &= \mathbf{R}_i^{(a)} \tilde{\mathbf{y}}_i^{(a)} \\ &= \sqrt{\rho_i} \mathbf{R}_i^{(a)} \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} \mathbf{d}_i^{(b)} + \mathbf{R}_i^{(a)} \tilde{\mathbf{n}}_i^{(a)}.\end{aligned}\quad (9)$$

C. Deriving the MSE of the received data stream

We can now formulate the MSE of the SU $i^{(a)}$. Using (9), the MSE matrix of the SU $i^{(a)}$ can be written as

$$\begin{aligned}\text{MSE}_i^{(a)} &= \mathbb{E} \left\{ \left(\hat{\mathbf{d}}_i^{(b)} - \mathbf{d}_i^{(b)} \right) \left(\hat{\mathbf{d}}_i^{(b)} - \mathbf{d}_i^{(b)} \right)^H \right\} \\ &= \left(\sqrt{\rho_i} \mathbf{R}_i^{(a)} \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} - \mathbf{I}_{d_i} \right) \left(\sqrt{\rho_i} \mathbf{R}_i^{(a)} \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} - \mathbf{I}_{d_i} \right)^H \\ &+ \mathbf{R}_i^{(a)} \Sigma_i^{(a)} \left(\mathbf{R}_i^{(a)} \right)^H.\end{aligned}\quad (10)$$

As mentioned before, the SUs are located within the service range of L PUs, for which the SUs should provide protection according to a QoS-based criterion. We assume that the PUs are equipped with N receive antennas. The received interference signal at the l th PU from SU $i^{(b)}$ is expressed as

$$\mathbf{z}_{i,l}^{(b)} = \sqrt{\mu_{i,l}^{(b)}} \mathbf{G}_{i,l}^{(b)} (\mathbf{x}_i^{(b)} + \mathbf{c}_i^{(b)}), \quad i \in \mathcal{K}, \quad b = 1, 2, \quad l \in \mathcal{L}, \quad (11)$$

where $\mathbf{G}_{i,l}^{(b)} \in \mathbb{C}^{N \times N_i}$ is the channel between the l th PU and $i^{(b)}$ th SU, which is modeled similar to $\mathbf{H}_{ij}^{(ab)}$ discussed in Section II-A, and $\mu_{i,l}^{(b)}$ is the average power gain of $\mathbf{G}_{i,l}^{(b)}$. Using (11), the power of the interference resulting from the $i^{(b)}$ th SU at the l th PU can be written as

$$\begin{aligned}I_{i,l}^{(b)} \left(\mathbf{V}_i^{(b)} \right) &= \mu_{i,l}^{(b)} \text{tr} \left\{ \mathbf{G}_{i,l}^{(b)} \left(\mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \right. \right. \\ &\left. \left. + \kappa \text{diag} \left(\mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \right) \right) \left(\mathbf{G}_{i,l}^{(b)} \right)^H \right\}.\end{aligned}\quad (12)$$

III. SUM-MSE MINIMIZATION

We take sum-MSE as the performance measure to design the transceivers under a transmit power constraint imposed to the SUs and an interference power constraint at the l th PU, which can be formulated as follows

$$\min_{\mathbf{V}, \mathbf{R}} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{MSE}_i^{(a)} \right\} \quad (13a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (13b)$$

$$I_{i,l}^{(b)} \left(\mathbf{V}_i^{(b)} \right) \leq \lambda_{i,l}^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad l \in \mathcal{L}, \quad (13c)$$

where $P_i^{(b)}$ is the power constraint at the $i^{(b)}$ th SU transmitter and $\lambda_{i,l}^{(b)}$ is the maximum average interference power that can be generated by SU $i^{(b)}$ and is received by the l th PU receiver [2], [3], [4], and $\mathbf{V}(\mathbf{R}) = \left\{ \mathbf{V}_i^{(b)} \left(\mathbf{R}_i^{(b)} \right) : \forall (i, b) \right\}$ is the set of all transmitting (receiving) beamforming matrices.

Fixing the transmit beamforming matrix, the optimal receive beamforming matrices at the SU $i^{(a)}$ is the MMSE receive filter which can be expressed as

$$\begin{aligned} \mathbf{R}_i^{(a)*} &= \arg \min_{\mathbf{R}_i^{(a)}} \text{tr} \left\{ \mathbf{MSE}_i^{(a)} \right\} \quad (14) \\ &= \sqrt{\rho_i} \left(\mathbf{V}_i^{(b)} \right)^H \left(\mathbf{H}_{ii}^{(ab)} \right)^H \\ &\quad \times \left(\rho_i \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \left(\mathbf{H}_{ii}^{(ab)} \right)^H + \boldsymbol{\Sigma}_i^{(a)} \right)^{-1}. \end{aligned}$$

Substituting (14) to the objective function $\mathbf{MSE}_i^{(a)}$ in (13a) gives us $\mathbf{C}_i^{(a)}(\mathbf{V})$, so-called error matrix for the node $i^{(a)}$ given that the MMSE receive filter is applied, and it can be written as

$$\begin{aligned} \mathbf{C}_i^{(a)}(\mathbf{V}) &= \mathbf{I}_{d_i} - \rho_i \left(\mathbf{V}_i^{(b)} \right)^H \left(\mathbf{H}_{ii}^{(ab)} \right)^H \quad (15) \\ &\quad \times \left(\rho_i \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \left(\mathbf{H}_{ii}^{(ab)} \right)^H + \boldsymbol{\Sigma}_i^{(a)} \right)^{-1} \\ &\quad \times \mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)}. \end{aligned}$$

Substituting $\mathbf{C}_i^{(a)}(\mathbf{V})$ into the objective function (13a), and writing $\mathbf{Q}_i^{(b)} = \mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H$, the problem of determining the

optimum transmit beamforming matrices under fixed receiver matrices can be rewritten as³

$$\max_{\mathbf{Q}} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \quad (16a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (16b)$$

$$I_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} \right) \leq \lambda_{i,l}^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad l \in \mathcal{L}, \quad (16c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (16d)$$

where $\mathbf{Q} = \left\{ \mathbf{Q}_i^{(b)} : \forall (i, b) \right\}$, and the matrix $\mathbf{A}_i^{(a)}(\mathbf{Q})$ is defined as

$$\begin{aligned} \mathbf{A}_i^{(a)}(\mathbf{Q}) &= \rho_i \mathbf{H}_{ii}^{(ab)} \mathbf{Q}_i^{(b)} \left(\mathbf{H}_{ii}^{(ab)} \right)^H \\ &\quad \times \left(\rho_i \mathbf{H}_{ii}^{(ab)} \mathbf{Q}_i^{(b)} \left(\mathbf{H}_{ii}^{(ab)} \right)^H + \tilde{\boldsymbol{\Sigma}}_i^{(a)} \right)^{-1}. \quad (17) \end{aligned}$$

Here, $\tilde{\boldsymbol{\Sigma}}_i^{(a)}$ in (17) and $I_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} \right)$ in (16c) are obtained by replacing $\mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H$ in (8) and (12) with $\mathbf{Q}_i^{(b)}$, respectively.

A. Imperfect CSI Model

According to the system model, the channels between the secondary transmitters and the primary receivers are assumed to be imperfectly known at the secondary side. The imperfect CSI is modeled using the deterministic norm-bounded error model [1]-[3], which is expressed as

$$\mathbf{G}_{i,l}^{(b)} \in \mathcal{G}_{i,l}^{(b)} = \left\{ \tilde{\mathbf{G}}_{i,l}^{(b)} + \boldsymbol{\Lambda}_{i,l}^{(b)} : \|\boldsymbol{\Lambda}_{i,l}^{(b)}\|_F \leq \theta_{i,l}^{(b)} \right\}, \quad \forall (i, b, l).$$

In the above equation, $\tilde{\mathbf{G}}_{i,l}^{(b)}$, $\boldsymbol{\Lambda}_{i,l}^{(b)}$ and $\theta_{i,l}^{(b)}$ denote the nominal value of the CSI, the error matrix, and the uncertainty bounds, respectively.

³Since the first term, i.e., the identity matrix in (15) has no effect in the optimization problem, we only consider the second term in (15), and the negative sign in front of the second term in (15) changes the minimization problem (13) to a maximization problem (16).

$$\begin{aligned} \boldsymbol{\Sigma}_i^{(a)} &\approx \rho_i \kappa \mathbf{H}_{ii}^{(ab)} \text{diag} \left(\mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \right) \left(\mathbf{H}_{ii}^{(ab)} \right)^H + \eta_{ii} \kappa \mathbf{H}_{ii}^{(aa)} \text{diag} \left(\mathbf{V}_i^{(a)} \left(\mathbf{V}_i^{(a)} \right)^H \right) \left(\mathbf{H}_{ii}^{(aa)} \right)^H \\ &\quad + \beta \rho_i \text{diag} \left(\mathbf{H}_{ii}^{(ab)} \mathbf{V}_i^{(b)} \left(\mathbf{V}_i^{(b)} \right)^H \left(\mathbf{H}_{ii}^{(ab)} \right)^H \right) + \beta \eta_{ii} \text{diag} \left(\mathbf{H}_{ii}^{(aa)} \mathbf{V}_i^{(a)} \left(\mathbf{V}_i^{(a)} \right)^H \left(\mathbf{H}_{ii}^{(aa)} \right)^H \right) \\ &\quad + \sum_{j \neq i}^K \sum_{c=1}^2 \eta_{ij}^{(ac)} \left[\mathbf{H}_{ij}^{(ac)} \left(\mathbf{V}_j^{(c)} \left(\mathbf{V}_j^{(c)} \right)^H + \kappa \text{diag} \left(\mathbf{V}_j^{(c)} \left(\mathbf{V}_j^{(c)} \right)^H \right) \right) \left(\mathbf{H}_{ij}^{(ac)} \right)^H \right] \\ &\quad + \sum_{j \neq i}^K \sum_{c=1}^2 \beta \eta_{ij}^{(ac)} \text{diag} \left(\mathbf{H}_{ij}^{(ac)} \mathbf{V}_j^{(c)} \left(\mathbf{V}_j^{(c)} \right)^H \left(\mathbf{H}_{ij}^{(ac)} \right)^H \right) + \mathbf{I}_{M_i}. \quad (8) \end{aligned}$$

With the imperfect CSI, the optimization problem in (16) can be rewritten as

$$\max_{\mathbf{Q}} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \quad (18a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (18b)$$

$$I_{i,l}^{(b)}(\mathbf{Q}_i^{(b)}) \leq \lambda_{i,l}^{(b)}, \quad \forall \mathbf{G}_{i,l}^{(b)} \in \mathcal{G}_{i,l}^{(b)}, \quad \forall (i, b, l), \quad (18c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad i \in \mathcal{K}, \quad b = 1, 2. \quad (18d)$$

IV. ROBUST TRANSCEIVER DESIGN

Solving the optimization problem (18) is difficult, since the objective function is non-convex. In order to solve it, we first linearize the objective function in (18a) and then employ a block-coordinate ascent solver. This can be accomplished as follows: for fixed matrices $\hat{\mathbf{Q}}_i^{(b)} = \left\{ \mathbf{Q}_j^{(c)} : \forall (j, c) \neq (i, b) \right\}$, we linearize the convex function $f_i^{(a)}(\cdot)$ in (19) around a feasible point $\tilde{\mathbf{Q}}_i^{(b)}$, and thus (locally) approximate the objective function (18a) as detailed below:

$$\begin{aligned} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} &= \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \\ &+ \underbrace{\sum_{(j,c) \neq (i,a)} \text{tr} \left\{ \mathbf{A}_j^{(c)}(\mathbf{Q}) \right\}}_{f_i^{(a)}(\mathbf{Q}_i^{(b)}, \hat{\mathbf{Q}}_i^{(b)})} \\ &\approx \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} + f_i^{(a)}(\tilde{\mathbf{Q}}_i^{(b)}, \hat{\mathbf{Q}}_i^{(b)}) \\ &+ \text{tr} \left\{ \left(\mathbf{D}_i^{(b)} \right)^H \left(\mathbf{Q}_i^{(b)} - \tilde{\mathbf{Q}}_i^{(b)} \right) \right\} \end{aligned} \quad (19)$$

Here $\mathbf{D}_i^{(b)} = \left. \frac{\partial f_i^{(a)}(\mathbf{Q}_i^{(b)}, \hat{\mathbf{Q}}_i^{(b)})}{\partial \mathbf{Q}_i^{(b)}} \right|_{\mathbf{Q}_i^{(b)} = \tilde{\mathbf{Q}}_i^{(b)}}$. One can derive the latter expression as follows:

$$\begin{aligned} \mathbf{D}_i^{(b)} &= -\eta_{ii} \left[\kappa \text{diag} \left(\left(\mathbf{H}_{ii}^{(bb)} \right)^H \left(\mathbf{Y}_i^{(b)} \right)^{-1} \mathbf{X}_i^{(b)} \left(\mathbf{Y}_i^{(b)} \right)^{-1} \mathbf{H}_{ii}^{(bb)} \right) \right. \\ &+ \beta \left(\mathbf{H}_{ii}^{(bb)} \right)^H \text{diag} \left(\left(\mathbf{Y}_i^{(b)} \right)^{-1} \mathbf{X}_i^{(b)} \left(\mathbf{Y}_i^{(b)} \right)^{-1} \right) \mathbf{H}_{ii}^{(bb)} \\ &+ \sum_{j \neq i} \sum_{c=1}^2 -\eta_{ji}^{(cb)} \left[\left(\mathbf{H}_{ji}^{(cb)} \right)^H \left(\mathbf{Y}_j^{(c)} \right)^{-1} \mathbf{X}_j^{(c)} \left(\mathbf{Y}_j^{(c)} \right)^{-1} \mathbf{H}_{ji}^{(cb)} \right. \\ &+ \kappa \text{diag} \left(\left(\mathbf{H}_{ji}^{(cb)} \right)^H \left(\mathbf{Y}_j^{(c)} \right)^{-1} \mathbf{X}_j^{(c)} \left(\mathbf{Y}_j^{(c)} \right)^{-1} \mathbf{H}_{ji}^{(cb)} \right) \\ &\left. + \beta \left(\mathbf{H}_{ji}^{(cb)} \right)^H \text{diag} \left(\left(\mathbf{Y}_j^{(c)} \right)^{-1} \mathbf{X}_j^{(c)} \left(\mathbf{Y}_j^{(c)} \right)^{-1} \right) \mathbf{H}_{ji}^{(cb)} \right], \end{aligned} \quad (20)$$

where $\mathbf{X}_i^{(b)} = \rho_i \mathbf{H}_{ii}^{(ba)} \mathbf{Q}_i^{(a)} \left(\mathbf{H}_{ii}^{(ba)} \right)^H$, and $\mathbf{Y}_i^{(a)} = \mathbf{X}_i^{(a)} + \tilde{\Sigma}_i^{(a)}$. Using the approximation in (19), the problem of com-

puting $\mathbf{Q}_i^{(b)}$ under fixed $\hat{\mathbf{Q}}_i^{(b)}$ is expressed as

$$\max_{\mathbf{Q}_i^{(b)}} \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} + \text{tr} \left\{ \left(\mathbf{D}_i^{(b)} \right)^H \mathbf{Q}_i^{(b)} \right\} \quad (21a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad (21b)$$

$$I_{i,l}^{(b)}(\mathbf{Q}_i^{(b)}) \leq \lambda_{i,l}^{(b)}, \quad \forall \mathbf{G}_{i,l}^{(b)} \in \mathcal{G}_{i,l}^{(b)}, \quad (21c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}. \quad (21d)$$

Since $f_i^{(a)}(\cdot)$ is a convex function of $\mathbf{Q}_i^{(b)}$, the local linear approximation using Taylor expansion is a lower bound which is tight at the current point $\tilde{\mathbf{Q}}_i^{(b)}$. This approximation enables the optimization problem to be solved efficiently through a numerical iterative algorithm. In fact, the idea of iteratively optimizing lower bounds of a non-convex function has been used in the literature for different contexts [14], [45]-[48]. The main idea is that the linearizing point in each iteration is selected such that the iterative algorithm gets closer to optimal point in every iteration. Roughly speaking, the main idea of this iterative method is similar to the gradient based methods. In the first iteration, we start with an arbitrary point and continue at every iteration by replacing the original objective function by its approximation obtained at the previous iteration until the utility improvement is less than a given tolerance.

In what follows next, similar to [3] we focus on the constraint (21c), in an attempt to transform the problem (21) into an equivalent problem that is easier to solve.

A. Reformulating the interference constraint

With the aim of further simplifying the problem (21), the following proposition will be proven helpful in order to relax the semi-infiniteness of interference constraint (21c).

Proposition 1: There exists $\epsilon_i^{(b)} \geq 0$, so that the interference constraint in (21c) can be written in linear-matrix-inequality form as

$$\begin{bmatrix} \mathbf{B}_i^{(b),[1]} & \mathbf{B}_i^{(b),[2]} \\ \left(\mathbf{B}_i^{(b),[2]} \right)^H & \mathbf{B}_i^{(b),[3]} \end{bmatrix} \succeq \mathbf{0}, \quad (22)$$

where

$$\mathbf{B}_{i,l}^{(b),[1]} = \epsilon_{i,l}^{(b)} \mathbf{I}_{N_i N} - \left[\mathbf{I}_N \otimes \left(\mathbf{Q}_i^{(b)} + \kappa \text{diag} \left(\mathbf{Q}_i^{(b)} \right) \right) \right], \quad (23)$$

$$\begin{aligned} \mathbf{B}_{i,l}^{(b),[2]} &= - \left[\text{vec} \left(\mathbf{Q}_i^{(b)} \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right) \right. \\ &\quad \left. + \kappa \text{vec} \left(\text{diag} \left(\mathbf{Q}_i^{(b)} \right) \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right) \right], \end{aligned} \quad (24)$$

$$\begin{aligned} \mathbf{B}_{i,l}^{(b),[3]} &= \frac{\lambda_{i,l}^{(b)}}{\mu_{i,l}^{(b)}} - \epsilon_{i,l}^{(b)} \left(\theta_{i,l}^{(b)} \right)^2 \\ &- \text{tr} \left\{ \tilde{\mathbf{G}}_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} + \kappa \text{diag} \left(\mathbf{Q}_i^{(b)} \right) \right) \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right\}. \end{aligned} \quad (25)$$

Proof: Applying the properties $\text{tr} \left\{ \mathbf{Z}^H \mathbf{A} \mathbf{Z} \right\} = \text{vec}(\mathbf{Z})^H (\mathbf{I} \otimes \mathbf{A}) \text{vec}(\mathbf{Z})$, $\text{tr} \left\{ \mathbf{B}^H \mathbf{Z} \right\} = \text{vec}(\mathbf{B})^H \text{vec}(\mathbf{Z})$ from [49] on

the interference constraint (21c), we obtain the inequality given in (26) at the bottom of the page. We further use the lemma given below from [50] to relax the semi-infiniteness of the constraint (26).

Lemma 1: Given $N \times N$ Hermitian matrices \mathbf{D} , \mathbf{A} ; $N \times 1$ vector \mathbf{b} , and the scalars c, e , there exists an $\bar{\mathbf{x}}$ satisfying $\bar{\mathbf{x}}^H \mathbf{D} \bar{\mathbf{x}} < e$. Then the inequality

$$\mathbf{x}^H \mathbf{A} \mathbf{x} + 2\Re \{ \mathbf{b}^H \mathbf{x} \} + c \geq 0, \quad \forall \mathbf{x}^H \mathbf{D} \mathbf{x} \leq e$$

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

$$\begin{bmatrix} \epsilon \mathbf{D} + \mathbf{A} & \mathbf{b} \\ \mathbf{b}^H & c - e\epsilon \end{bmatrix} \succeq \mathbf{0}.$$

By mapping \mathbf{D} , \mathbf{A} , \mathbf{b} , c and e in Lemma 1 with the ones given in (26), the interference constraint (21c) can be equivalently reformulated as in Proposition 1. \blacksquare

B. Resulting optimization problem and algorithm

Using the matrix inversion lemma on $\mathbf{A}_i^{(a)}$ in (17), and substituting it to the objective function (21a), the resulting optimization problem of computing $\mathbf{Q}_i^{(b)}$ under fixed $\tilde{\mathbf{Q}}_i^{(b)}$ can be expressed as

$$\min_{\mathbf{Q}_i^{(b)}, \epsilon_{i,l}^{(b)}} \quad \text{tr} \left\{ \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2} \left(\mathbf{Y}_i^{(a)} \right)^{-1} \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2} \right\} - \text{tr} \left\{ \left(\mathbf{D}_i^{(b)} \right)^H \mathbf{Q}_i^{(b)} \right\} \quad (27a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad (27b)$$

$$\begin{bmatrix} \mathbf{B}_{i,l}^{(b),[1]} & \mathbf{B}_{i,l}^{(b),[2]} \\ \left(\mathbf{B}_{i,l}^{(b),[2]} \right)^H & \mathbf{B}_{i,l}^{(b),[3]} \end{bmatrix} \succeq \mathbf{0}, \quad (27c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad \epsilon_{i,l}^{(b)} \geq 0, \quad (27d)$$

where $\mathbf{Y}_i^{(a)}$ is defined in (20).

We now introduce an auxiliary variable \mathbf{Y} such that $\mathbf{Y} \succeq \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2} \left(\mathbf{Y}_i^{(a)} \right)^{-1} \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2}$, and use Schur complement.

The resulting problem is expressed as

$$\min_{\mathbf{Q}_i^{(b)}, \epsilon_{i,l}^{(b)}, \mathbf{Y}} \quad \text{tr} \{ \mathbf{Y} \} - \text{tr} \left\{ \left(\mathbf{D}_i^{(b)} \right)^H \mathbf{Q}_i^{(b)} \right\} \quad (28a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad (28b)$$

$$\begin{bmatrix} \mathbf{B}_{i,l}^{(b),[1]} & \mathbf{B}_{i,l}^{(b),[2]} \\ \left(\mathbf{B}_{i,l}^{(b),[2]} \right)^H & \mathbf{B}_{i,l}^{(b),[3]} \end{bmatrix} \succeq \mathbf{0}, \quad (28c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad \epsilon_{i,l}^{(b)} \geq 0, \quad (28d)$$

$$\begin{bmatrix} \mathbf{Y}_i^{(a)} & \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2} \\ \left(\tilde{\Sigma}_i^{(a)} \right)^{1/2} & \mathbf{Y} \end{bmatrix} \succeq \mathbf{0}. \quad (28e)$$

The problem (28) is a SDP, which can be efficiently solved in polynomial time by standard interior point methods. The proposed robust algorithm for the sum-MSE optimization problem (28) that uses the SDP method is given in Algorithm 1.

Algorithm 1 Robust Sum-MSE minimization.

- 1: Initialize $\mathbf{Q}_i^{(b)} = \mathbf{0}$, $i = 1, \dots, K$, $b = 1, 2$.
 - 2: **repeat**
 - 3: **for** $i = 1 : K$ **do**
 - 4: **for** $b = 1 : 2$ **do**
 - 5: Compute $\mathbf{D}_i^{(b)}$ from (20).
 - 6: Update $\mathbf{Q}_i^{(b)}$ by solving (28).
 - 7: **end for**
 - 8: **end for**
 - 9: **until** convergence of the objective function (16a) or maximum number of iterations is reached.
 - 10: Compute the receive filter $\mathbf{R}_i^{(a)}$, $\forall (i, a)$ from (14).
-

C. Discussion

1) *Convergence:* Since (18a) is bounded from above, by employing the monotone convergence theorem we can conclude that the proposed algorithm converges, and as shown in Appendix, it converges to a stationary point.

$$\begin{aligned} & -\text{vec} \left(\left(\mathbf{\Lambda}_{i,l}^{(b)} \right)^H \right)^H \underbrace{\left[\mathbf{I}_N \otimes \left(\mathbf{Q}_i^{(b)} + \kappa \text{diag} \left(\mathbf{Q}_i^{(b)} \right) \right) \right]}_{-\mathbf{A}} \underbrace{\text{vec} \left(\left(\mathbf{\Lambda}_{i,l}^{(b)} \right)^H \right)}_{\mathbf{x}} \\ & - 2\Re \left\{ \underbrace{\left[\text{vec} \left(\mathbf{Q}_i^{(b)} \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right)^H + \kappa \text{vec} \left(\text{diag} \left(\mathbf{Q}_i^{(b)} \right) \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right)^H \right]}_{-\mathbf{b}^H} \text{vec} \left(\left(\mathbf{\Lambda}_{i,l}^{(b)} \right)^H \right) \right\} \\ & + \underbrace{\frac{\lambda_{i,l}^{(b)}}{\mu_{i,l}^{(b)}} - \text{tr} \left\{ \tilde{\mathbf{G}}_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} + \kappa \text{diag} \left(\mathbf{Q}_i^{(b)} \right) \right) \left(\tilde{\mathbf{G}}_{i,l}^{(b)} \right)^H \right\}}_c \geq 0, \quad \text{vec} \left(\mathbf{\Lambda}_{i,l}^{(b)} \right)^H \underbrace{\mathbf{I}_{NN_i}}_{\mathbf{D}} \text{vec} \left(\mathbf{\Lambda}_{i,l}^{(b)} \right) \leq \underbrace{\left(\theta_{i,l}^{(b)} \right)^2}_e. \end{aligned} \quad (26)$$

2) *Channel Acquisition*: Channel estimation between SUs can be accomplished using standard signal processing techniques via training through pilots and feedback [5], and can be assumed as perfect, which has been widely used in the CR literature. Channel estimation between SUs and PUs is more challenging, because PUs are unlikely to cooperate with SUs. If the primary system adopts the TDD scheme, by exploiting channel reciprocity, the channels between SUs and PUs can be acquired at the SUs by overhearing the transmissions between PU transmitter and receiver pair [5]-[10]. If applying a TDD scheme is not feasible, blind beamforming techniques can be employed [5]. Other methods to acquire the CSI knowledge between PUs and SUs is 1) through environmental learning [11], [12], 2) by exchange of CSI between the PUs and SUs through a *band manager*, which mediates between the two parties [6], [7], [13], and 3) the primary system can cooperate with the secondary system to exchange the channel estimates [5]. Of course, since the primary and secondary systems are not fully coordinated, the quality of these channel estimates will be degraded. Hence, we choose to model these imperfections by considering norm-bounded estimation errors for the links between the secondary transmitters and primary receivers. Note that after SUs obtain the CSIs of the channels, they report them to the central scheduler to perform resource allocation/transceiver design in each time slot [8].

3) *Implementation*: The proposed algorithm requires the existence of a central scheduler at the secondary system, which collects all channel matrices, and then computes and distributes the beamforming matrices of all SUs. In particular, the central scheduler collects all the SUs channels $\mathbf{H}_{ij}^{(ab)}$, all the estimated SU-PU channels $\tilde{\mathbf{G}}_i^{(b)}$ and all the confidence intervals $\theta_i^{(b)}$. In an interference channel (or an ad-hoc network), the scheduler can reside at any node in the network. In a dynamic environment, the scheduler can be adaptively elected among the eligible nodes in the network [52]-[53]. The election can be done based on the capacity of a node, the status of a node, or its location.

It should be noted that the proposed algorithm can also be implemented in a distributed fashion, i.e., at the n th iteration, each SU node $i^{(b)}$ can update its $\mathbf{Q}_i^{(b),[n]}$ locally through the problem (28) provided that the received interference-plus-noise covariance matrix $\tilde{\Sigma}_i^{(a),[n]}$ is locally estimated through measurements [54]-[55], and the matrices $\mathbf{Y}_j^{(c),[n]}$ and $\mathbf{X}_j^{(c),[n]}$ are obtained from the neighboring SUs via local message passing. Note that the matrices $\mathbf{Y}_j^{(c),[n]}$, $\mathbf{X}_j^{(c),[n]}$ and $\mathbf{Q}_i^{(b),[n-1]}$ obtained from the previous iteration are used to compute the gradient in (20). The overall distributed scheme is described in Algorithm 2. Note that the distributed scheme requires each SU to collect only local channel information, and thus improves the scalability.

Algorithm 2 Distributed Robust Sum-MSE minimization.

```

1: Initialize  $\mathbf{Q}_i^{(b)} = \mathbf{0}$ ,  $i = 1, \dots, K$ ,  $b = 1, 2$ .
2: repeat ( $n = 1, 2, \dots$ )
3:   for  $i = 1 : K$  do
4:     for  $b = 1 : 2$  do
5:       Node  $i^{(b)}$  acquires the channels  $\mathbf{H}_{ji}^{(cb)}$ ,  $\forall (j, c) \neq (i, b)$  from its neighbors.
6:       Transmit  $\mathbf{Y}_i^{(a),[n]}$  and  $\mathbf{X}_i^{(a),[n]}$  to neighboring nodes.
7:       Receive  $\mathbf{Y}_j^{(c),[n]}$  and  $\mathbf{X}_j^{(c),[n]}$ ,  $(j, c) \neq (i, a)$  from neighboring nodes.
8:       Compute  $\mathbf{D}_i^{(b),[n]}$  from (20).
9:       Measure  $\tilde{\Sigma}_i^{(a),[n]}$ .
10:      Update  $\mathbf{Q}_i^{(b),[n]}$  by solving (28).
11:      Compute the receive filter  $\mathbf{R}_i^{(a)}$  from (14).
12:    end for
13:  end for
14: until convergence of the objective function (21a) or maximum number of iterations is reached.

```

V. EXTENSIONS

A. Total-Interference Constraint

When the individual interference power constraints $\lambda_{i,l}^{(b)}$ are not available, it is practical to consider the total-interference power constraint that reflects the aggregate interference received by each PU, due to secondary transmissions [14]. 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 [4]. In this paper, we will consider deterministic interference constraints as assumed in [3], [14]. Particularly, we assume that the PU imposing the interference constraint, has already computed its maximum tolerable interference threshold. The problem (16) is modified under the total-interference power constraint as follows

$$\max_{\mathbf{Q}} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \quad (29a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (29b)$$

$$\sum_{i=1}^K \sum_{b=1}^2 I_{i,l}^{(b)}(\mathbf{Q}_i^{(b)}) \leq \lambda_l, \quad l \in \mathcal{L}, \quad (29c)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (29d)$$

where λ_l is the maximum total interference power constraint at the l th PU.

Unlike an off-line algorithm, in an on-line algorithm, the intermediate results must be also feasible. Dual decomposition method [56] used for the coupled constraints is not suitable for our purpose, because the problem (29) is non-convex and non-separable, and thus the duality gap is generally non-zero. This results in unfeasible intermediate results while running the on-line algorithm, which can violate the interference constraint (29c). We will apply the primal decomposition technique [14] to tackle the coupled interference constraint (29c). Similar to [3], by introducing two sets of auxiliary variables,

$\psi_i^{(b)}$ and $\xi_i^{(b)}$, the problem (29) can be equivalently reformulated as

$$\max_{\mathbf{Q}, \psi, \xi} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \quad (30a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (30b)$$

$$I_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} \right) \leq \xi_{i,l}^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad l \in \mathcal{L}, \quad (30c)$$

$$0 \leq \xi_{i,l}^{(b)} \leq \psi_{i,l}^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad l \in \mathcal{L}, \quad (30d)$$

$$\sum_{i=1}^K \sum_{b=1}^2 \psi_{i,l}^{(b)} \leq \lambda_l, \quad l = 1, \dots, L, \quad (30e)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (30f)$$

where $\psi(\boldsymbol{\xi}) = \left\{ \psi_{i,l}^{(b)}(\xi_{i,l}^{(b)}) : \forall (i, b, l) \right\}$.

For fixed $\psi_{i,l}^{(b)}$, the inner maximization subproblem is written as

$$\mathcal{P}(\psi) := \max_{\mathbf{Q}, \xi} \sum_{i=1}^K \sum_{a=1}^2 \text{tr} \left\{ \mathbf{A}_i^{(a)}(\mathbf{Q}) \right\} \quad (31a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{Q}_i^{(b)} \right\} \leq P_i^{(b)}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (31b)$$

$$I_{i,l}^{(b)} \left(\mathbf{Q}_i^{(b)} \right) \leq \xi_{i,l}^{(b)}, \quad \forall (i, b, l), \quad (31c)$$

$$0 \leq \xi_{i,l}^{(b)} \leq \psi_{i,l}^{(b)}, \quad \forall (i, b, l), \quad (31d)$$

$$\mathbf{Q}_i^{(b)} \succeq \mathbf{0}, \quad i \in \mathcal{K}, \quad b = 1, 2, \quad (31e)$$

which is similar to the problem in (18), and can be solved similarly using the local linearization method discussed in Section IV. After solving the problem (31) under fixed ψ , the per-SU transmitter interference constraint $\psi_{i,l}^{(b)}$ is updated by solving the following master problem, which allocates the total interference constraint λ among SUs:

$$\max_{\psi} \quad \mathcal{P}(\psi) \quad (32a)$$

$$\text{s.t.} \quad \sum_{i=1}^K \sum_{b=1}^2 \psi_{i,l}^{(b)} \leq \lambda_l, \quad l \in \mathcal{L}, \quad (32b)$$

$$\psi_{i,l}^{(b)} \geq 0, \quad i \in \mathcal{K}, \quad b = 1, 2. \quad (32c)$$

Since the local linearization of the problem (31) is convex, the master problem can be solved by applying the subgradient projection method. The subgradient of $\mathcal{P}(\psi)$ with respect to $\psi_{i,l}^{(b)}$ is the optimal Lagrange multiplier $\tau_{i,l}^{(b)}$ corresponding to the constraint (31d) [56, Chap. 5]. At the n th iteration, the subgradient projection updating the interference constraints is expressed as

$$\psi[n+1] = \text{Proj}[\psi[n] + s[n]\boldsymbol{\tau}[n]], \quad (33)$$

where $\boldsymbol{\tau} = \left\{ \tau_{i,l}^{(b)} : \forall (i, b, l) \right\}$, s is a positive step size, and $\text{Proj}[\cdot]$ denotes projection onto the convex feasible set defined by (32b)-(32c).

After the maximization problem (31) is solved, each SU transmits the local scalar Lagrange multiplier $\tau_{i,l}^{(b)}[n]$ to a central SU node for the update of $\psi[n+1]$ using (33) and then the central node will distribute these interference constraints back to the SUs.

B. Full-Duplex Cognitive Cellular Systems

In this subsection, we show that the algorithm proposed for the MIMO FD CR system also holds for FD cognitive cellular systems, in which a FD BS communicates with K UL and J DL HD mode users, simultaneously as seen in Fig. 2. The BS is equipped with M_0 and N_0 transmit and receive antennas, respectively. The number of antennas of the k -th UL user and the j -th DL user are denoted by M_k and N_j , respectively. The number of data streams transmitted from the k -th UL user (to the j -th DL user) is denoted by d_k^{UL} (d_j^{DL}).

$\mathbf{H}_k^{UL} \in \mathbb{C}^{N_0 \times M_k}$ and $\mathbf{H}_j^{DL} \in \mathbb{C}^{N_j \times M_0}$ represent the k -th UL channel and the j -th DL channel, respectively. $\mathbf{H}_0 \in \mathbb{C}^{N_0 \times M_0}$ is the self-interference channel from the transmit antennas of the BS to the receive antennas of the BS. $\mathbf{H}_{jk}^{DU} \in \mathbb{C}^{N_j \times M_k}$ denotes the CCI channel from the k -th UL user to the j -th DL user.

The vector of source symbols transmitted by the k -th UL user is denoted as \mathbf{s}_k^{UL} . It is assumed that the symbols are i.i.d. with unit power, i.e., $\mathbb{E}[\mathbf{s}_k^{UL} (\mathbf{s}_k^{UL})^H] = \mathbf{I}_{d_k^{UL}}$. Similarly, the transmit symbol vector for the j -th DL user is denoted by \mathbf{s}_j^{DL} , with $\mathbb{E}[\mathbf{s}_j^{DL} (\mathbf{s}_j^{DL})^H] = \mathbf{I}_{d_j^{DL}}$. Denoting the precoders for the data streams of the k -th UL and j -th DL user as $\mathbf{V}_k^{UL} \in \mathbb{C}^{M_k \times d_k^{UL}}$, and $\mathbf{V}_j^{DL} \in \mathbb{C}^{M_0 \times d_j^{DL}}$, respectively, the transmitted signal of the k -th UL user and that of the BS can be written, respectively, as

$$\mathbf{x}_k^{UL} = \mathbf{V}_k^{UL} \mathbf{s}_k^{UL}, \quad \mathbf{x}_0 = \sum_{j=1}^J \mathbf{V}_j^{DL} \mathbf{s}_j^{DL}. \quad (34)$$

The described multi-user MIMO system suffers from self-interference and CCI. The signal received by the BS and the one received by the j -th DL user can be written, respectively, as

$$\mathbf{y}_0 = \sum_{k=1}^K \mathbf{H}_k^{UL} (\mathbf{x}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{H}_0 (\mathbf{x}_0 + \mathbf{c}_0) + \mathbf{e}_0 + \mathbf{n}_0, \quad (35)$$

$$\mathbf{y}_j^{DL} = \mathbf{H}_j^{DL} (\mathbf{x}_0 + \mathbf{c}_0) + \sum_{k=1}^K \mathbf{H}_{jk}^{DU} (\mathbf{x}_k^{UL} + \mathbf{c}_k^{UL}) + \mathbf{e}_j^{DL} + \mathbf{n}_j^{DL}, \quad (36)$$

where $\mathbf{n}_0 \in \mathbb{C}^{N_0}$ and $\mathbf{n}_j^{DL} \in \mathbb{C}^{N_j}$ denote the AWGN vector with zero mean and identity covariance matrix at the the BS and the j -th DL user, respectively. In (35)-(36), \mathbf{c}_k^{UL} (\mathbf{c}_0) is the transmitter distortion at the k -th UL user (BS), which is modeled as in (4), and \mathbf{e}_j^{DL} (\mathbf{e}_0) is the receiver distortion at the j -th DL user (BS), which is modeled as in (5).

The received signals are processed by linear decoders, denoted as $\mathbf{U}_k^{UL} \in \mathbb{C}^{N_0 \times d_k^{UL}}$, and $\mathbf{U}_j^{DL} \in \mathbb{C}^{N_j \times d_j^{DL}}$ by the

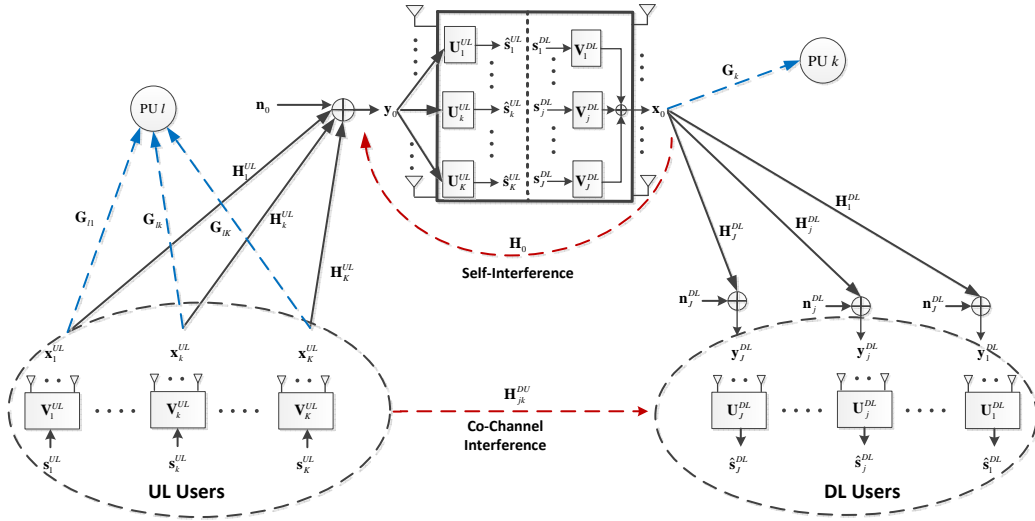


Fig. 2. Multi-user MIMO system with a full-duplex-capable BS.

BS receiver and the j -th DL user, respectively. As a result, the estimated data streams of the k -th UL user at the BS and j -th DL user are expressed as

$$\hat{\mathbf{s}}_k^{UL} = (\mathbf{U}_k^{UL})^H \mathbf{y}_0, \quad \hat{\mathbf{s}}_j^{DL} = (\mathbf{U}_j^{DL})^H \mathbf{y}_j^{DL}. \quad (37)$$

Using these estimates, the MSE of the k th UL and j th DL user can be written as in (38) and (39), respectively, shown at the bottom of the following page. Here, in (38), Σ_k^{UL} is the covariance matrix of the aggregate interference-plus-noise terms at the k -th UL, and can be approximated, under $\beta \ll 1$ and $\kappa \ll 1$, as in (40) at the bottom of the following page. The covariance matrix of the aggregate interference-plus-noise terms at the j -th DL user, Σ_j^{DL} in (39) can be defined similarly, by replacing \mathbf{H}_m^{UL} , \mathbf{V}_m^{UL} , and \mathbf{H}_0 in (40) with \mathbf{H}_{jm}^{DL} , \mathbf{V}_m^{DL} , and \mathbf{H}_{jm}^{DU} , respectively.

The power of the interference resulting from the k th UL user and the BS transmitter at the l th PU can be written, respectively as

$$I_{k,l}^{UL} = \text{tr} \left\{ \mathbf{G}_{lk} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H + \kappa \text{diag} \left(\mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H \right) \right) \mathbf{G}_{lk}^H \right\}, \quad (41)$$

$$I_l^{DL} = \sum_{j=1}^J \text{tr} \left\{ \mathbf{G}_l \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H + \kappa \text{diag} \left(\mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \right) \right) \mathbf{G}_l^H \right\}, \quad (42)$$

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

C. Joint Beamforming Design

The optimization problem can be formulated as:

$$\min_{\mathbf{U}_k^{UL}, \mathbf{U}_j^{DL}, \mathbf{V}_k^{UL}, \mathbf{V}_j^{DL}} \sum_{k=1}^K \text{tr} \left\{ \text{MSE}_k^{UL} \right\} + \sum_{j=1}^J \text{tr} \left\{ \text{MSE}_j^{DL} \right\} \quad (43a)$$

$$\text{s.t.} \quad \text{tr} \left\{ \mathbf{V}_k^{UL} (\mathbf{V}_k^{UL})^H \right\} \leq P_k, \quad k \in \mathcal{S}^{UL}, \quad (43b)$$

$$\sum_{j=1}^J \text{tr} \left\{ \mathbf{V}_j^{DL} (\mathbf{V}_j^{DL})^H \right\} \leq P_0, \quad (43c)$$

$$I_{k,l}^{UL} \leq \lambda_{k,l}^{UL}, \quad k \in \mathcal{S}^{UL}, \quad l \in \mathcal{L}, \quad (43d)$$

$$I_l^{DL} \leq \lambda_l^{DL}, \quad l \in \mathcal{L}, \quad (43e)$$

where P_k in (43b) is the transmit power constraint at the k -th UL user, and P_0 in (43c) is the total power constraint at the BS transmitter side. Moreover, $\lambda_{k,l}^{UL}$ and λ_l^{DL} are the maximum average interference power that can be generated by k th UL user and BS, respectively and is received by the l th PU receiver. We use \mathcal{S}^{UL} and \mathcal{S}^{DL} to represent the set of K UL and J DL channels, respectively.

1) *Simplification of Notations*: To simplify the notations, we will combine UL and DL channels, similar to [21]. Denoting \mathbf{H}_{ij} and \mathbf{n}_i as

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

$$\mathbf{G}_{lj} = \begin{cases} \mathbf{G}_{lj}, & j \in \mathcal{S}^{UL}, \\ \mathbf{G}_l, & j \in \mathcal{S}^{DL}, \end{cases}$$

and referring to \mathbf{V}_i^X , \mathbf{U}_i^X , Σ_i^X , $X \in \{UL, DL\}$ as \mathbf{V}_i , \mathbf{U}_i , Σ_i , respectively, the MSE of i -th link, $i \in \mathcal{S} \triangleq \mathcal{S}^{UL} \cup \mathcal{S}^{DL}$

can be written as

$$\mathbf{MSE}_i = (\mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i - \mathbf{I}) (\mathbf{U}_i^H \mathbf{H}_{ii} \mathbf{V}_i - \mathbf{I})^H + \mathbf{U}_i^H \boldsymbol{\Sigma}_i \mathbf{U}_i, \quad (44)$$

where

$$\begin{aligned} \boldsymbol{\Sigma}_i = & \sum_{j \in \mathcal{S}, j \neq i} \mathbf{H}_{ij} \mathbf{V}_j \mathbf{V}_j^H \mathbf{H}_{ij}^H + \kappa \sum_{j \in \mathcal{S}} \mathbf{H}_{ij} \text{diag}(\mathbf{V}_j \mathbf{V}_j^H) \mathbf{H}_{ij}^H \\ & + \beta \sum_{j \in \mathcal{S}} \text{diag}(\mathbf{H}_{ij} \mathbf{V}_j \mathbf{V}_j^H \mathbf{H}_{ij}^H) + \mathbf{I}, \end{aligned} \quad (45)$$

and the interference power generated by the j th UL and BS at the l -th PU in (41)-(42) can be rewritten as

$$I_{j,l} = \text{tr} \{ \mathbf{G}_{lj} (\mathbf{V}_j \mathbf{V}_j^H + \kappa \text{diag}(\mathbf{V}_j \mathbf{V}_j^H)) \mathbf{G}_{lj}^H \}, \quad (46)$$

$$I_l = \sum_{j \in \mathcal{S}^{DL}} \text{tr} \{ \mathbf{G}_{lj} (\mathbf{V}_j \mathbf{V}_j^H + \kappa \text{diag}(\mathbf{V}_j \mathbf{V}_j^H)) \mathbf{G}_{lj}^H \}. \quad (47)$$

2) *Transceiver Design*: Using the simplified notations, the optimization problem (43) can be rewritten as

$$\min_{\mathbf{V}_i, \mathbf{U}_i} \sum_{i \in \mathcal{S}} \text{tr} \{ \mathbf{MSE}_i \} \quad (48a)$$

$$\text{s.t.} \quad \text{tr} \{ \mathbf{V}_i \mathbf{V}_i^H \} \leq P_i, \quad i \in \mathcal{S}^{UL}, \quad (48b)$$

$$\sum_{i \in \mathcal{S}^{DL}} \text{tr} \{ \mathbf{V}_i \mathbf{V}_i^H \} \leq P_0, \quad (48c)$$

$$I_{j,l} \leq \lambda_{j,l}^{UL}, \quad j \in \mathcal{S}^{UL}, \quad l \in \mathcal{L}, \quad (48d)$$

$$I_l \leq \lambda_l^{DL}, \quad l \in \mathcal{L}. \quad (48e)$$

The optimization problem (48) has the same formulation as the optimization problem in (13), which focuses on the CR MIMO interference channels, and thus under the fixed receive beamforming matrices, we can apply the transceiver design algorithm proposed in Section IV, accordingly.

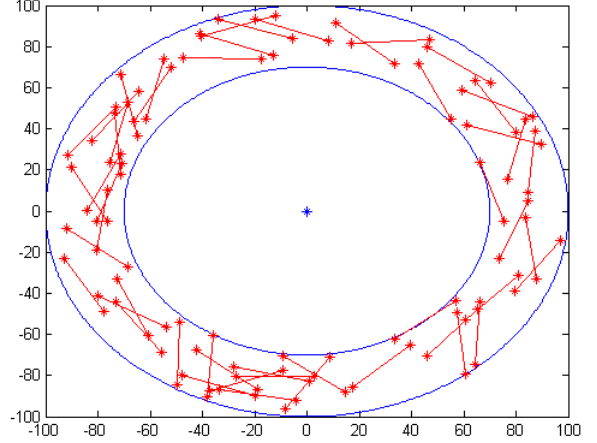


Fig. 3. Simulation Setup. The PU receiver (blue star) is located at the origin, the SU pairs (red connected stars) are located between 70 – 100m away from the PU receiver.

VI. SIMULATION RESULTS

In this section, we numerically investigate the proposed robust sum-MSE minimization algorithm using the general purpose convex optimization program CVX package designed for Matlab [57]. 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 100, and the results are averaged over 800 independent channel realizations.

The distance between the desired links is set to $d_i = 30\text{m}$. The PU receiver is located at a distance from the SUs that is uniformly distributed over 70 – 100m. Fig. 3 illustrates the examined system scenario, where a PU receiver is located at the origin, and the K pairs of SUs are located between 70 – 100m away from it. For brevity, we assume that the maximum transmit-power for all SUs is the same,

$$\mathbf{MSE}_k^{UL} = \left((\mathbf{U}_k^{UL})^H \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}} \right) \left((\mathbf{U}_k^{UL})^H \mathbf{H}_k^{UL} \mathbf{V}_k^{UL} - \mathbf{I}_{d_k^{UL}} \right)^H + (\mathbf{U}_k^{UL})^H \boldsymbol{\Sigma}_k^{UL} \mathbf{U}_k^{UL}, \quad (38)$$

$$\mathbf{MSE}_j^{DL} = \left((\mathbf{U}_j^{DL})^H \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}} \right) \left((\mathbf{U}_j^{DL})^H \mathbf{H}_j^{DL} \mathbf{V}_j^{DL} - \mathbf{I}_{d_j^{DL}} \right)^H + (\mathbf{U}_j^{DL})^H \boldsymbol{\Sigma}_j^{DL} \mathbf{U}_j^{DL}. \quad (39)$$

$$\begin{aligned} \boldsymbol{\Sigma}_k^{UL} = & \sum_{m \neq k}^K \mathbf{H}_m^{UL} \mathbf{V}_m^{UL} (\mathbf{V}_m^{UL})^H (\mathbf{H}_m^{UL})^H + \kappa \sum_{m=1}^K \mathbf{H}_m^{UL} \text{diag}(\mathbf{V}_m^{UL} (\mathbf{V}_m^{UL})^H) (\mathbf{H}_m^{UL})^H \\ & + \sum_{m=1}^J \mathbf{H}_0 (\mathbf{V}_m^{DL} (\mathbf{V}_m^{DL})^H + \kappa \text{diag}(\mathbf{V}_m^{DL} (\mathbf{V}_m^{DL})^H)) \mathbf{H}_0^H + \beta \sum_{m=1}^K \text{diag}(\mathbf{H}_m^{UL} \mathbf{V}_m^{UL} (\mathbf{V}_m^{UL})^H (\mathbf{H}_m^{UL})^H) \\ & + \beta \sum_{m=1}^J \text{diag}(\mathbf{H}_0 \mathbf{V}_m^{DL} (\mathbf{V}_m^{DL})^H \mathbf{H}_0^H) + \mathbf{I}_{N_0}. \end{aligned} \quad (40)$$

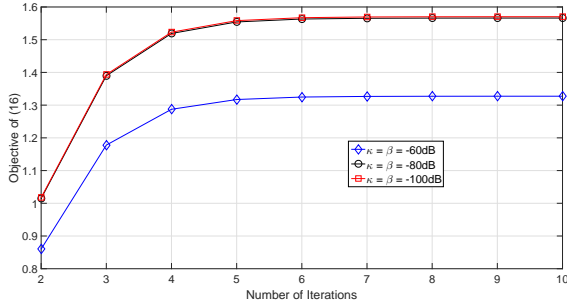


Fig. 4. Convergence of the objective function in (16a). Here, $K = 4$, $s = 0.2$, $N = 2$.

i.e., $P = P_i^{(b)}$, $\forall (i, b)$. The path loss obeys the model $d^{-\zeta}$, where d is the distance between nodes, and $\zeta = 3.5$ is the path-loss exponent. The maximum transmit-powers are set so that the (maximum) signal-to-noise ratio (SNR) is defined as $\text{SNR} = P d_i^{-\zeta} = 15\text{dB}$. The total interference threshold is set to $4 \times 10^{-7}\text{W}$, and for simplicity, it is equally split among the SUs. The channel uncertainty is set to $\theta_i^{(b)} = s \|\tilde{\mathbf{G}}_i^{(b)}\|_F$, with $s \in (0, 1]$ [2]. The transmitter/receiver distortion parameters are chosen as $\kappa = \beta = -60\text{dB}$. For brevity, we set the same number of transmit and receive antennas at each node, i.e., $M_i = N_i = N$, $i = 1, \dots, K$.⁴

Fig. 4 visualizes the evolution of Algorithm 1, i.e., the convergence of the objective function in (16a) under different distortion levels. The monotonic increase of the objective function (16a) can be verified, and is seen to converge in less than 10 iterations.

In our second example, we will compare our robust FD precoding scheme with the robust HD one for different values of the channel estimation errors, i.e., s parameter mentioned above. The performance is going to be measured by means of the average sum rate of the secondary system. Note that for a HD system, transmission is carried out in two time-slots, i.e., in the first time slot, all the SUs on the left hand side in Fig. 1 transmit to their peers on the right, whereas in the second time-slot, these roles are reversed. As a result, although self-interference does not exist, CCI is present and the sum-rate should be divided by 2 because of the two time-slots transmission. It can be seen from Fig. 5 that as the size of the uncertainty region increases, the performance of the FD system degrades more, and the performance gap between the considered FD and HD systems decreases. This degradation in performance of the FD system is explained as follows. Since there are more interference channels (self-interference and CCI) in FD systems, as the uncertainty level of the channels increases, the system performance of the FD system can become seriously degraded. This indicates that the channel

⁴Note that although the nodes in i th link have $N_i + M_i$ antennas in total, similar to [28], [43], we assume that only N_i (M_i) antennas can be used for transmission (reception) in HD mode. 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 each node in the i th link only has N_i transmission front-ends and M_i receiving front-ends, and does not carry out antenna partitioning.

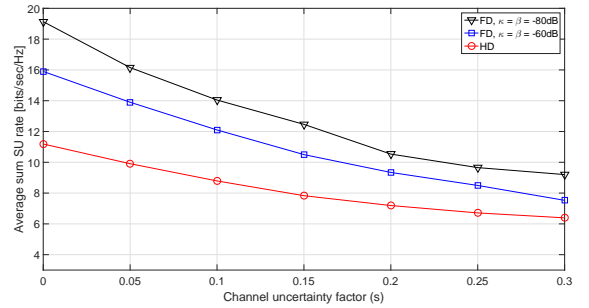


Fig. 5. Sum-Rate comparison of FD and HD systems with respect to channel uncertainty factor s . Here, $K = 2$, $N = 2$.

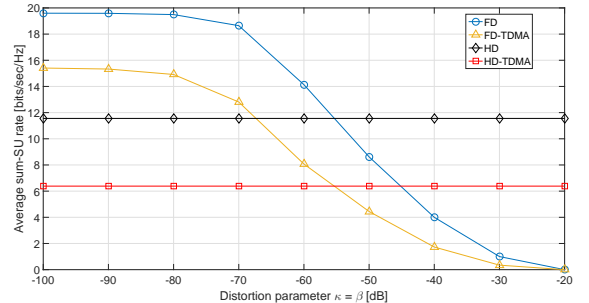


Fig. 6. Sum-Rate comparison of FD and HD systems with respect to transmitter/receiver distortion κ , β . Here, $K = 4$, $s = 0.2$, $N = 2$.

estimation is a critical factor for successful deployment of FD systems.

In our next example, we will compare the optimized FD system along with the optimized corresponding baseline systems, i.e., HD, HD-TDMA, and FD-TDMA systems in terms of sum-rate performance for different $\kappa = \beta$ values. In FD-TDMA, in the first time slot, only the first pair transmits and receives in FD mode. In the second time slot, only the second pair transmits and receives in FD mode, and in the K th time slot, only the K th pair transmits and receives in FD mode. In this case, the sum-rate should be divided by the number of time slots (or pairs), in our case K . In HD-TDMA, all

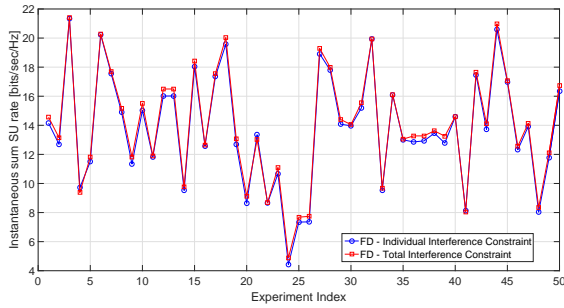


Fig. 7. Sum-Rate comparison of individual and total power constrained problems. Here, $K = 4$, $s = 0.2$, $\kappa = \beta = -60\text{dB}$, $N = 2$.

the nodes transmit sequentially, so we need $2K$ time slots. Therefore, we have neither self-interference nor CCI. As seen in Fig. 6, the performance of the HD and HD-TDMA systems are not affected by κ and β values, and at high self-interference cancellation levels, the FD systems achieves around 1.6 times more sum-rate than that of the corresponding HD system. It is also worth mentioning that the performance of FD system drops below that of HD scheme around $\kappa = \beta = -55\text{dB}$, and below that of HD-TDMA scheme around $\kappa = \beta = -45\text{dB}$.

This result is not in line with the ones obtained in [58]-[60]. In [58], [59], the authors tackle sum-power minimization and sum-rate maximization problems in FD MIMO interference channels, respectively, and it is concluded in [59] that HD mode surprisingly outperforms the FD mode in terms of achieved throughput. The difference between the results in [59] and our paper can be explained as below:

- Unlike [59], in our paper, we assume that an (imperfect) analog domain interference-cancellation or passive suppression has been implemented in the FD radios, and the main strong self-interference is canceled (please see (6)). However, due to the transmit and receive distortions, the residual self-interference still exists in the baseband characterized by κ and β , respectively. As seen in Fig. 6, depending on the distortion level, FD gain over HD systems is positive or negative.
- Moreover, unlike [59], where an *antenna-conserved* scenario is considered, in our paper we consider an *RF-chain conserved* scenario [60]. As shown in [60], a FD radio performance is inferior to that of an HD MIMO radio in the antenna conserved scenario; however, the spectral efficiency can be potentially doubled when the RF-chain conserved scenario is considered.

To compare the performances of the proposed individual and total power constrained algorithms discussed in Section IV and V-A, respectively, in Fig. 7, Fig. 7 depicts the achieved sum-rate in 50 different experiments. Since both algorithms converge to locally optimal solutions, there is no guarantee that one will always outperform the other. However, it is seen that the total-power constrained algorithm usually performs better than the individual power one because the total-power constraint is less conservative.

In Fig. 8, we compare the user-MSEs achieved by the

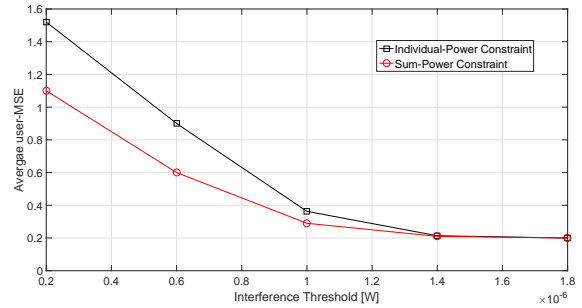


Fig. 8. MSE comparison of individual and total power constrained problems with respect to interference constraint. Here, $K = 4$, $s = 0.2$, $N = 2$, $\kappa = \beta = -60\text{dB}$.

TABLE I. AVERAGE RATE GAIN OF FD SYSTEM OVER HD SYSTEM

K	2	4	6	8	10
$\frac{\text{FD-HD}}{\text{HD}}$	29.86%	21.78%	15.38%	6.74%	0.45%

individual power and total-power constrained problems with respect to interference constraint. It is seen that as the interference constraint increases, the gap between two curves decreases. The reason is that when the interference constraint is small, the SUs transmit with low transmit-powers in order not to violate the PU interference constraint. On the other hand, when the interference constraint is high, the effect of interference constraint on the performance of the system diminishes, and thus two systems achieve the same MSE.

Finally, the gains achieved by the FD system over HD system is depicted in Table I. Here, $s = 0.2$, $\kappa = \beta = -60\text{dB}$, $N = 2$. It is seen that as the number of pairs increases the FD gain decreases, since the number of CCI terms introduced with the FD mode increases, and the performance of the FD system deteriorates.

VII. CONCLUSION

In this work, we have studied the robust MSE-based transceiver design problem for a FD cognitive MIMO interference channel that suffers from self-interference and CCI under the limited DR at the transmitters and receivers, as well as from norm-bounded channel uncertainties. Since the globally optimal solution is difficult to obtain due to the non-convex nature of the problem, a SDP-based cyclic block coordinate ascent algorithm shown to converge to a stationary point is proposed. Simulation results have confirmed that the sum-rate achieved by the FD system is higher than that of HD system under reasonable self-interference cancellation levels. Moreover, the importance of accurate channel estimation for FD systems is emphasized via numerical results.

APPENDIX

The proof is based on the convergence of the block coordinate descent method in [3], [51], which shows that every limit point of the algorithm satisfies the first-order optimality conditions over the Cartesian product of the closed convex sets. Let $\bar{\mathbf{Q}} = (\bar{\mathbf{Q}}_1^{(1)}, \bar{\mathbf{Q}}_1^{(2)}, \dots, \bar{\mathbf{Q}}_K^{(1)}, \bar{\mathbf{Q}}_K^{(2)})$ be a limit point of the sequence $\mathbf{Q}^{[n]}$ which represents the set of all transmit-covariance

matrices at iteration n . Moreover, let $\{\mathbf{Q}^{[n_j]} \mid j = 1, 2, \dots\}$ be a subsequence that converges to $\bar{\mathbf{Q}}$. As proved in [51], $\lim_{j \rightarrow \infty} \mathbf{Q}_i^{(b), [n_j+1]} = \bar{\mathbf{Q}}_i^{(b)}$, $\forall (i, b)$.

Let us denote $\mathcal{U}(\cdot)$ and $\mathcal{U}_i^{(b)}(\cdot)$ as the objective functions in (18a) and (21a), respectively. Since $\mathbf{Q}_i^{(b), [n_j+1]}$ is the local (and also global) maximum of $\mathcal{U}_i^{(b)}(\mathbf{Q}_i^{(b), [n_j+1]}, \hat{\mathbf{Q}}_i^{(b), [n_j]})$, we have

$$\Re \left\{ \text{tr} \left\{ \nabla_i^b \mathcal{U}_i^{(b)}(\mathbf{Q}_i^{(b), [n_j+1]}, \hat{\mathbf{Q}}_i^{(b), [n_j]})^H \times \left(\mathbf{Q}_i^{(b)} - \mathbf{Q}_i^{(b), [n_j+1]} \right) \right\} \right\} \leq 0, \quad \forall \mathbf{Q}_i^{(b)} \in \mathcal{C}_i^{(b)}, \quad (49)$$

where $\nabla_i^b \mathcal{U}_i^{(b)}(\cdot)$ denotes the gradient of $\mathcal{U}_i^{(b)}(\cdot)$ with respect to $\mathbf{Q}_i^{(b)}$, and $\mathcal{C}_i^{(b)}$ is the feasible set for the problem (21), i.e., $\mathcal{C}_i^{(b)} = \left\{ \mathbf{Q}_i^{(b)} : \mathbf{Q}_i^{(b)} \in (21b) - (21d) \right\}$. Taking the limit as $j \rightarrow \infty$, and using the fact that $\nabla_i^b \mathcal{U}_i^{(b)}(\bar{\mathbf{Q}}) = \nabla_i^b \mathcal{U}(\bar{\mathbf{Q}})$, it is easy to show that

$$\Re \left\{ \text{tr} \left\{ \nabla_i^b \mathcal{U}(\bar{\mathbf{Q}})^H \times \left(\mathbf{Q}_i^{(b)} - \bar{\mathbf{Q}}_i^{(b)} \right) \right\} \right\} \leq 0, \quad \forall \mathbf{Q}_i^{(b)} \in \mathcal{C}_i^{(b)}, \quad (50)$$

which establishes the stationarity of $\bar{\mathbf{Q}}$ and completes the proof.

REFERENCES

- [1] 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.
- [2] 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.
- [3] Y. Zhang, E. Dall'Anese, 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] L. Zhang, Y.-C. Liang, and Y. Xin, "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 38-51, Jan. 2008.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] J. M. Peha, "Approaches to spectrum sharing," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 10-12, Feb. 2005.
- [14] 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.
- [15] Z. Zhang, X. Chai, K. Long, A. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: Self-interference cancellation, protocol design, and relay selection," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 128-137, May 2015.
- [16] I. Chih-Lin et al., "Towards green and soft: A 5G perspective," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 66-73, Feb. 2014.
- [17] 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.
- [18] 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, March 2015.
- [19] D. Bharadia and S. Katti, "Full duplex MIMO radios," *USENIX NSDI*, pp. 359-372, 2014.
- [20] 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, Sep. 2014.
- [21] S. Li, R. Murch, and V. Lau, "Linear transceiver design for full-duplex multi-user MIMO system," *IEEE Int. Conf. Commun. (ICC)*, pp. 4921-4926, June 2014.
- [22] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of loopback self-interference in full-duplex MIMO relays," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5983-5993, Dec. 2011.
- [23] H. A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P. J. Smith, "Low complexity end-to-end performance optimization in MIMO full-duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 13, pp. 913-927, Jan. 2014.
- [24] 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.
- [25] A. C. Cirik, O. Taghizadeh, R. Mathar and T. Ratnarajah, "QoS considerations for full duplex multi-user MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, Feb. 2016.
- [26] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-aho "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 801-815, Mar. 2015.
- [27] A. C. Cirik, "Fairness considerations for full-duplex multi-user MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 4, no. 4, pp. 361-364, Aug. 2015.
- [28] 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.
- [29] Y. Liao, L. Song, Z. Han, and Y. Li, "Full-duplex cognitive radio: A new design paradigm for enhancing spectrum usage," *IEEE Commun. Mag.*, vol. 53, no. 5, pp. 138-145, May 2015.
- [30] Y. Liao, T. Wang, L. Song, and Z. Han, "Listen-and-Talk: Full-duplex cognitive radio networks," *IEEE Global Commun. Conf (Globecom)*, pp. 3068-3073, Dec. 2014.
- [31] J. Heo, H. Ju, S. Park, E. Kim, and D. Hong, "Simultaneous sensing and transmission in cognitive radio," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1948-1959, Apr. 2014.

- [32] H. Kim, S. Lim, H. Wang, and D. Hong, "Optimal power allocation and outage analysis for cognitive full duplex relay systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3754-3765, Oct. 2012.
- [33] G. Zheng, I. Krikidis, and B. Ottersten, "Full-duplex cooperative cognitive radio with transmit imperfections," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2498-2511, May 2013.
- [34] Y. Deng, K. J. Kim, T. Q. Duong, M. ElKashlan, G. K. Karagiannidis, and A. Nallanathan, "Full-duplex spectrum sharing in cooperative single carrier systems," *IEEE Wireless Communications and Networking Conf. (WCNC'15)*, pp. 25-30, Mar. 2015.
- [35] 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, June 2015.
- [36] J. G. D. Forney, "Shannon meets Wiener II: On MMSE estimation in successive decoding schemes," in *Proc. Allerton Conf. Commun., Control, and Computing*, Monticello, IL, Sep. 2004.
- [37] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2381-2401, Sep. 2003.
- [38] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 698-713, Feb. 2009.
- [39] Y. Rong, X. Tang, and Y. Hua, "A unified framework for optimizing linear nonregenerative multicarrier MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 57, no. 12, pp. 4837-4851, Dec. 2009.
- [40] 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.
- [41] 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.
- [42] A. C. Cirik, Y. Rong, and Y. Hua, "Achievable rates of full-duplex MIMO radios in fast fading channels with imperfect channel estimation," *IEEE Trans. Signal Processing*, vol. 62, no. 15, pp. 3874-3886, Aug. 2014.
- [43] 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, Sep. 2014.
- [44] 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, June 2013.
- [45] C. T. K. Ng and H. Huang, "Linear precoding in cooperative MIMO cellular networks with limited coordination clusters," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1446-1454, Dec. 2010.
- [46] L.-N. Tran, "An iterative precoder design for successive zero-forcing precoded systems," *IEEE Commun. Lett.*, vol. 16, no. 1, pp. 16-18, Jan. 2012.
- [47] D. Nguyen, L. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038-4050, Aug. 2013.
- [48] A. Khabbazi-basmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of DC programming problems," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5478-5493, Oct. 2012.
- [49] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. <http://matrixcookbook.com> [online], Nov. 2012.
- [50] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [51] M. Razaviyayn, M. Sanjabi, and Z. Q. Luo, "Linear transceiver design for interference alignment: Complexity and computation," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 2896-2910, May 2012.
- [52] Y. Rong and Y. Hua, "Optimal power schedule for distributed MIMO links," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 2896-2900, Aug. 2008.
- [53] Y. Rong and Y. Hua, "Space-time power scheduling of MIMO links – Fairness and QoS considerations," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 2, pp. 171-180, Apr. 2008.
- [54] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331-4340, Sep. 2011.
- [55] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "A distributed numerical approach to interference alignment and applications to wireless interference networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3309-3322, Jun. 2011.
- [56] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA: Athena Scientific, 1999.
- [57] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming*, 2008 [Online]. Available: <http://stanford.edu/boyd/cvx>.
- [58] D. N. Nguyen, M. Krunz, "Be responsible: A novel communications scheme for full-duplex MIMO radios," *IEEE INFOCOM Conf.*, pp. 1733-1741, May 2015.
- [59] D. N. Nguyen, M. Krunz, and S. Hanly, "On the throughput of full-duplex MIMO in the multi-link case," *IEEE WiOpt Conf.*, pp. 421-428, May 2014.
- [60] S. Barghi, A. Khojastepour, and K. Sundaresan, "Characterizing the throughput gain of single cell MIMO wireless systems with full duplex radios," *IEEE WiOpt*, pp. 68-74, May 2012.



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, 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, 5G non-orthogonal multiple-access (NOMA), MIMO signal processing, and convex optimization.



Miltiades C. Filippou (S'12-M'15) was born in Athens, Greece in 1984. He received his Dipl. Eng. degree in Electrical and Computer Engineering from the National Technical University of Athens (NTUA), Greece, in 2007. From November 2007 until July 2009 he was occupied at the Wireless Communications laboratory of the same school as a post-graduate researcher in the area of satellite communications. In July 2014, he received his Ph.D. degree in Electronics and Telecommunications from Telecom ParisTech, France. Between September 2014 and October 2015, he was a Research Fellow at the Institute for Digital Communications (IDCOM), University of Edinburgh, U.K. Since November 2015 he works as a Standards and Research Engineer at Intel Deutschland GmbH, Neubiberg, Germany. Dr. Filippou is a member of the IEEE and of the Technical Chamber of Greece. His current research interests include: 5G air interface design, millimeter-wave and cooperative communications, as well as spectrum sharing systems.



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. His research interests include signal processing and information theoretic aspects of 5G 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 300 publications in these areas and holds four U.S. patents. He is currently the coordinator of the FP7 projects ADEL (3.7M€) in the area of licensed shared access for 5G wireless networks. Previously, he was the coordinator of the FP7 project 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., and an associate editor of the IEEE Transactions on Signal Processing.