Full-duplex MIMO Radios: A Greener Networking Solution

Diep N. Nguyen[1], Marwan Krunz[1,2], and Eryk Dutkiewicz[1]
[1]Faculty of Engineering and Information Technology, University of Technology Sydney,
[2]Department of Electrical and Computer Engineering, University of Arizona

***Abstract*—Relative to half-duplex (HD) radios, in-band Full-duplex (FD) radios have the potential to double *a link's capacity*. However, such gain may not necessarily extend to the *network-wide throughput*, which may actually degrade under FD radios due to excessive network interference. This paper identifies the unique advantages of FD radios and leverages multi-input multi-output (MIMO) communications to translate the FD spectral efficiency gain at the PHY level to the throughput and power efficiency gain at the network layer. We first derive sufficient conditions under which FD-MIMO radios can *asymptotically* double the throughput of the same network of HD-MIMO ones. Specifically, if a network of $2N$ HD radios (N links) can achieve a total throughput of dN bps (i.e., d bps per link), then an FD-capable network with the same number of links and network/channel realization can achieve $2N(d-1)$ bps (i.e., $(d-1)$ bps per direction of a bidirectional link). To leverage this theoretical gain, we exploit the "spatial signature" readily captured in the network interference to design a MAC protocol that allows multiple FD links to concurrently communicate while adapting their radiation patterns to minimize network interference. The protocol does not require any feedback nor coordination among nodes. Extensive simulations show that the proposed MAC design dramatically outperforms traditional CSMA-based and the non-orthogonal multiple access (NOMA) protocols with either HD or FD radios w.r.t. both throughput and energy efficiency. Note that in the literature, network interference is often treated as colored noise that then gets whiten during the signal detection process. However, through our MAC protocol, we emphasize that, unlike random noise, network interference has its own structure that can be "mined" for "intelligence" to better align the transceiver's signal.**

***Index Terms*—Full-duplex, MIMO, Nash equilibrium, green communications, energy efficiency, capacity, networking, MAC layer.**

I. Introduction

The United Nations Climate Change Conference (COP21) has set an ambitious target to keep global warming below 2 degrees Celsius. To that end, every nation has committed to cut its annual greenhouse gas emission, e.g., carbon dioxide, to as low as zero before 2030. Such a target strongly motivates all industries, including wireless systems (only its clouds generate about 27 megatonnes of Carbon dioxide, equal to the environmental impact caused by 4.9 million new cars [2]) to look for greener alternatives. In this paper, we demonstrate that recent advances in self-interference suppression (SIS) at the wireless PHY layer can help to significantly conserve transmission power/energy. SIS allows a wireless device to transmit and receive simultaneously, i.e., perform full-duplex (FD) communications, on the same frequency [3] [4] and even using the same antenna array. Over the last few years, various SIS techniques have been demonstrated, including antenna cancelation, analog RF, and digital cancelation (see [5] and references therein). Latest developments have successfully suppressed self-interference down to the noise floor for both single [6] and multi-antenna (i.e., MIMO) [7] devices.

The spectral efficiency of an FD link has been shown to be nearly double that of a conventional half-duplex (HD) link [6] [7]. However, it has been reported (e.g., [8] [9] [10] [11]) that the network-wide (multiple links) throughput gain under FD radios is unexpectedly marginal or even negative compared to HD-based systems for both ad hoc and cellular setups. Most previous works on SIS focused on the throughput gain, with few exceptions (e.g., [12] [13] [14]), which investigated the energy/power efficiency of FD systems. This article attempts to translate the FD spectral efficiency gain at the *link level* into both throughput and energy/power efficiency gains at the *network level*.

Unlike HD radios, both ends of an FD link transmit at the same time. Thus, a set of mutually interfering FD links will now experience higher network interference, and subsequently, reduction in the spatial reuse. Although previous works (e.g., [8] [9] [10] [15]) identified the roots of throughput reduction in a network of FD radios, they didn't rigorously answer the question whether FD network throughput can *ever* double that of a HD network and how the FD network's power efficiency compares to that of an HD network. If throughput doubling is possible, then under what conditions? Seeking an answer to this question is critical in designing efficient MAC protocols for FD-based multi-user systems.

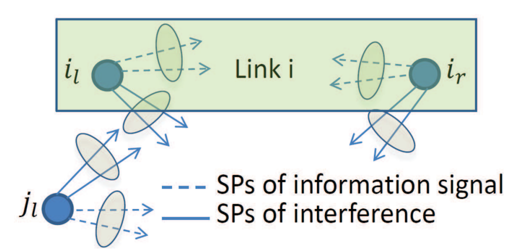


Fig. 1. Bidirectional link i is comprised of two FD MIMO radios i_l and i_r. From the product of precoding and channel gain matrices of an interfering FD MIMO radio j_l (captured at the RF chains of i_r), i_r can infer the SPs of information signal of link j (implicitly embedded in the precoding matrix of j_l) and SPs of interference induced on j_l from i_r (via channel reciprocity). i_l then can configure its radiation pattern to reduce its interference on j_r.

The SIS capability not only improves the spectral efficiency but also allows a wireless device to *instantaneously* discern the

This work has been supported in part by the Australian Research Council (Discovery Early Career Researcher Award DE150101092), and in part by US NSF (grants IIP-1265960, CNS-1513649, CNS-1563655, and CNS- 1731164). Preliminary results in this paper were presented at the IEEE INFOCOM Conference, Hong Kong, 2015 [1].

state of the medium while transmitting and instantly adjust its transmission strategy. This was leveraged in [16] to combat hidden/exposed terminals in CSMA-based protocols or and in [17] [18] to improve the spectrum sensing/awareness of an opportunistic access systems. For MIMO communications, the *interference/signal perceived by an FD MIMO radio provides it with much more valuable information other than just the busy/idle status of the medium*: It helps a node partially infer the "spatial signatures" (SPs) [19] [20] of the information signals intended to other nodes as well as the interference from the underlying node onto other nodes. That can be referred to as the *"deep sensing"* capability of an FD MIMO radio.

Recall that in MIMO communications, signal alignment can be realized via *precoding*, in which the information symbol vector is pre-multiplied with a precoding matrix before being placed on a Tx antenna array. A precoding matrix, or precoder, is a matrix of complex elements whose phase and amplitude can be tuned to control the radiation directions/beams of the signal [19]. SP of the mth data (or interfering) stream is the mth column vector of the corresponding channel gain matrix that describes the spatial direction along which the stream's received power is maximized. Note that for FD radios that use the same antenna array to transmit and receive simultaneously [7], channel reciprocity holds, i.e., the channel gain matrix on one direction is the transpose of that of the other direction. Hence, as shown in Fig. 1, the interference seen at the RF chains of radio i_r (product of channel gain matrix from j_l to i_r and j_l's Tx precoder) contains the SPs of link j's information signal (implicitly embedded in j_l's Tx precoder that is used to align j_l's signal along link j's SPs) *and* the SPs of the interference induced on j_l from i_r (via channel reciprocity).

Learning the above SPs together with the ability to *adjust the radiation pattern instantaneously* by tuning the phase and the amplitude of elements of its precoding matrix allow a FD MIMO node to minimize its required Tx power and reduce its interference on others. This ability does not exist in HD radios (where Tx and Rx have to take turns) nor single-antenna radios, which cannot control their radiation beams. Note that previous works (e.g., [9] [10] [16]) considered a protocol model (e.g., at most one single active link in a collision area) that does not allow links to coexist, and hence ignored the above advantage of FD MIMO radios, which in fact facilitates concurrent FD transmissions.

Although there have been recent reports on FD MIMO, e.g., [21], only MIMO beamforming has been considered. Note that for multi-antenna systems, precoding is a generalized form of beamforming (beamforming is precoding with a single data stream, i.e., precoder of rank one, e.g., [21]). In our work, we consider a general precoder with no constraint on its matrix rank. Various recent works, e.g., [22] [23] reported the energy inefficiency of FD systems, compared with HD one and even suggested using FD radios for short distances only (as in micro- or pico-cells) due to their excessive interference. In this work, we present a precoding design for FD MIMO that achieves much higher energy efficiency than HD radios by exploiting the spatial signatures in the network interference.

To establish the conditions that guarantee the superiority of FD over HD radios in a network setting, we consider the transmit power minimization problem for the entire network subject to rate constraints. Note that there is a large literature on sum-rate maximization subject to power constraints, e.g., [8] [9] [10]. Considering the power minimization problem allows us to derive sufficient conditions under which a set of rate demands can be met. We can then identify sufficient conditions under which the FD radio network can *asymptotically* double the throughput of the same network but with HD radios. Note that due to interference, the network-wide transmit power minimization problem is nonconvex. Hence, even with the availability of global network information, solving such a problem is prohibitively expensive. Seminal approaches (e.g., [20] [24] [25]) that studied the power minimization problem subject to SINR requirements for single-antenna HD radios are inapplicable to our FD MIMO setup that involves matrix operations. These matrix operations prevent us from obtaining a closed-form expression/projection (e.g., the well-known standard interference function in [24]) of precoders in terms of SINR/rate requirements.

Given the above, we formulate a noncooperative game in which FD links are players who aim to meet their rate demands by optimizing their precoders. To exploit captured SPs, instead of simply minimizing the transmit powers as in HD radios, an FD radio minimizes the sum of transmit powers on its antennas, weighted by the transpose of its interference covariance matrix (*as locally perceived by the radio*). Following our approach in [26], we use recession analysis [27] and the variational inequality theory [28] to provide sufficient conditions under which a Nash Equilibrium (NE) exists and a set of rates can be met. We also prove that the NE is unique. At this NE, if a network of $2N$ HD radios (N links) can achieve a total throughput of dN bps (i.e., d bps per link), then an FD-capable network with the same number of links and network/channel realization can achieve $2N(d-1)$ bps (i.e., $(d-1)$ bps per direction of a bidirectional link).

Extensive simulations show that for a given set of rate demands, the proposed approach is much more power-efficient than when HD radios are used or when FDs do not exploit SPs. The total power consumption under our approach is significantly less than N times that of the CSMA-based approach (where only one link is allowed to use the medium at a time) while the network throughput is N times higher. We also observe that the game converges quickly to its NE, facilitating the design of a practical MAC protocol, which we refer to as FD-MAC. As a performance benchmark, we use the augmented Lagrangian method to develop a centralized algorithm for the FD network-wide power minimization problem. Our main contributions can be summarized as follows:

- We establish sufficient conditions under which FD radios can double the network throughput.
- We identify and exploit the unique advantages (i.e., "deep sensing") of FD MIMO radios to enable the coexistence of multiple links, leading to significant energy/power savings. This is done by having FD MIMO radios instantly discern the medium at a finer scale (i.e., spatial signatures of other radios) and instantaneously adjust/adapt their radiation beams.
- We design an efficient MAC protocol for a network of FD radios. The proposed FD-MAC protocol does not require any feedback from or coordination between links,

as precoders are designed using only local information. Via simulations, FD-MAC is shown to achieve almost the same performance as its centralized (yet locally optimal) version (which aims to minimize the total network Tx power). FD-MAC yields much higher energy efficiency and throughout gain than traditional CSMA-based and the non-orthogonal multiple access (NOMA) protocols with either HD or FD radios. We extend the above two results to the multi-carrier scenario.

- We prove the existence of a unique NE to which the FD-MAC protocol converges. Simulations show that the FD-MAC converges to this NE fairly fast.

Throughout the paper, we use $(.)^*$ to denote the conjugate of a matrix, $(.)^H$ for its Hermitian transpose, tr$(.)$ for its trace, $|.|$ for the determinant, and $(.)^T$ for the matrix transpose. diag$_m(.)$ indicates the diagonal element (m, m) of a matrix, and sum$(.)$ gives the summation of all elements of a vector. $||.||$ denotes the Euclidean norm. Matrices and vectors are bold-faced.

The rest of the paper is as follows. In Section II, we present the network model and problem formulation. Conditions for the existence and uniqueness of the NE and rate-demand satisfaction, optimal precoders, and the MAC protocol are presented in Section III. The centralized algorithm is developed in IV. Numerical results are discussed in Section V, followed by concluding remarks in Section VI.

II. Network Model

Consider an ad hoc network of N FD-MIMO bidirectional links. The two ends of each link i, $i \in \mathcal{N} \stackrel{\text{def}}{=} \{1, ..., N\}$, operate simultaneously as a transmitter and a receiver. For simplicity, we denote the two FD radios/nodes of link i by the left radio i_l and the right radio i_r. Without loss of generality, each node is equipped with M antennas (our analysis is still applicable when nodes have different numbers of antennas). Let $\mathbf{H}_{ii}^{rl}$ ($\mathbf{H}_{ii}^{lr}$) denote the $M \times M$ channel gain matrix of the left-to-right (right-to-left) direction of link i. Due to channel reciprocity, $\mathbf{H}_{ii}^{rl} = (\mathbf{H}_{ii}^{lr})^T$. Each element of $\mathbf{H}_{ii}^{rl}$ is the multiplication of a distance- and frequency-dependent attenuation term and a random term that reflects multi-path fading (complex Gaussian variables with zero mean and unit variance). Let $\mathbf{H}_{ij}^{ll}$ and $\mathbf{H}_{ij}^{lr}$ denote the $M \times M$ interfering channel matrices from radios j_l and j_r of link j to radio i_l of link i, respectively, for any $i, j \in \mathcal{N}, i \neq j$. $\mathbf{H}_{ij}^{rl}$ is defined similarly.

Note that latest advances in SIS (e.g., [7] [5] and therein references) are able to suppress self-interference to the noise floor level (achieving 110dB or above SIS). However, to be pragmatic, we still account for the imperfect SIS by having $\mathbf{H}_{ii}^{ll}$ and $\mathbf{H}_{ii}^{rr}$ denote the self-interference channel matrices at radio i_l and i_r, respectively. g_{sis} denotes the self-interference suppression level (the ratio of residual suppressed signal to the transmit signal). Let $\mathbf{G}_i^l$ and $\mathbf{G}_i^r$ be the transmit precoding matrices at i_l and i_r, respectively. Let $\mathbf{x}_i^r$ denote the vector of transmit information symbols being placed on the antennas of radio i_r (for the right-to-left transmission of link i). The received signal vector $\mathbf{y}_i^l$ at the antennas of radio i_l is:

$$\mathbf{y}_i^l = \mathbf{H}_{ii}^{lr}\mathbf{G}_i^r\mathbf{x}_i^r + \sqrt{g_{sis}}\mathbf{H}_{ii}^{ll}\mathbf{G}_i^l\mathbf{x}_i^l + \sum_{j=1|j\neq i}^{N}\left(\mathbf{H}_{ij}^{ll}\mathbf{G}_j^l\mathbf{x}_j^l + \mathbf{H}_{ij}^{lr}\mathbf{G}_j^r\mathbf{x}_j^r\right) + \mathbf{N_o} \tag{1}$$

where the first term is the intended signal, the second term is the self-interference induced by the transmit chains of radio i_l, the third and fourth terms represent interference from the left and right radios of link j, and $\mathbf{N_o}$ is an $M \times 1$ complex Gaussian noise vector with identity covariance matrix $\mathbf{I}$, representing the normalized noise floor. The noise-plus-interference covariance matrix at radio i_r, $\mathbf{Q}_i^r$, is:

$$\mathbf{Q}_i^r = \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r{\mathbf{G}_i^r}^H{\mathbf{H}_{ii}^{rr}}^H + \sum_{j=1|j\neq i}^{N}\left(\mathbf{H}_{ij}^{rl}\mathbf{G}_j^l{\mathbf{G}_j^l}^H{\mathbf{H}_{ij}^{rl}}^H + \mathbf{H}_{ij}^{rr}\mathbf{G}_j^r{\mathbf{G}_j^r}^H{\mathbf{H}_{ij}^{rr}}^H\right).$$

Let c_i^l (c_i^r) denote the achieved throughput at node i_l (i_r) of link i. $\mathbf{G} \stackrel{\text{def}}{=} [\mathbf{G}_1^l \times \mathbf{G}_1^r \ldots \times \mathbf{G}_N^l \times \mathbf{G}_N^r]$ denotes the set of precoders from all radios and $\mathbf{G}_{-i_r}$ is the set of all precoders except that of i_r radio. Treating interference from other radios as colored noise, we have:

$$\begin{aligned} c_i^l(\mathbf{G}) &= \log|\mathbf{I} + {\mathbf{G}_i^r}^H{\mathbf{H}_{ii}^{lr}}^H[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r| \\ c_i^r(\mathbf{G}) &= \log|\mathbf{I} + {\mathbf{G}_i^l}^H{\mathbf{H}_{ii}^{rl}}^H[\mathbf{Q}_i^r]^{-1}\mathbf{H}_{ii}^{rl}\mathbf{G}_i^l| \end{aligned} \tag{2}$$

The network-wide power minimization problem subject to rate demands d_i^l (to be granted to receiver i_l) and d_i^r (to be granted to receiver i_r) is stated as follows:

$$\begin{aligned} &\underset{\{\mathbf{G_i^l}, \mathbf{G_i^r}, \forall i\}}{\text{minimize}} \sum_{i=1}^{N}\{\text{tr}(\mathbf{G}_i^l\mathbf{G}_i^{lH}) + \text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH})\} \\ &\text{s.t.} \quad \text{C1: } d_i^l \leq c_i^l, \ \ \forall i \in \mathcal{N} \\ &\qquad\ \ \text{C2: } d_i^r \leq c_i^r, \ \ \forall i \in \mathcal{N} \end{aligned} \tag{3}$$

In the above optimization, each radio is practically subject to a power constraint. However, to study the feasibility of *any given set of rate demands* (based solely on the network interference) regardless of the power budget constraint, we ignore the transmit power constraint. Seminal works (e.g., [24] [20] [25]), that studied the power minimization problem subject to rate demands, also neglected these constraints. Imposing a power budget would put a constraint on the set of rate demands that would have be achievable given the network geometry and interference. In terms of the optimization methodology, the power constraint at each node is an inequality with trace matrix operation that is convex w.r.t. the precoding matrix. As such, the following solution using convex optimization is still applicable.

III. Noncooperative Game Formulation

A. Formulation

The network-wide power minimization problem (3) is not convex, and hence is computationally expensive to solve even in a centralized manner. Additionally, collecting the required network information to solve (3) often involves excessive overhead. Existing works on HD-based systems formulate strategic noncooperative games to suboptimally solve such optimization in a distributed manner where the players are transmitting nodes. The transmit precoder $\tilde{\mathbf{G}}_i^r$ of radio i_r is found from:

$$\begin{aligned} &\underset{\{\mathbf{G}_i^r\}}{\text{minimize}} \;\; \text{tr}(\mathbf{G}_i^r \mathbf{G}_i^{rH}) \\ &\text{s.t.} \qquad d_i^l \leq c_i^l. \end{aligned} \tag{4}$$

Similarly, the transmit precoder $\mathbf{G}_i^l$ of radio i_l can be found by solving:

$$\begin{aligned} &\underset{\{\mathbf{G}_i^l\}}{\text{minimize}} \;\; \text{tr}(\mathbf{G}_i^l \mathbf{G}_i^{lH}) \\ &\text{s.t.} \qquad d_i^r \leq c_i^r. \end{aligned}$$

As mentioned above, an FD radio i_r can use its receive chain to gauge how much interference its antennas induce on others. Specifically, consider the transpose of the covariance matrix of interference-plus-noise perceived by radio i_r:

$$\begin{aligned} [\mathbf{Q}_i^r]^T =& \mathbf{I} + g_{sis}(\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r {\mathbf{G}_i^r}^H {\mathbf{H}_{ii}^{rr}}^H)^T \\ &+ \sum_{j=1|j\neq i}^{N} \left(({\mathbf{G}_j^l}^T \mathbf{H}_{ji}^{lr})^H ({\mathbf{G}_j^l}^T \mathbf{H}_{ji}^{lr}) + ({\mathbf{G}_j^r}^T \mathbf{H}_{ji}^{rr})^H ({\mathbf{G}_j^r}^T \mathbf{H}_{ji}^{rr}) \right) \end{aligned}$$

Let

$$\begin{aligned} \mathbf{S}_i^r \overset{\text{def}}{=}& g_{sis}(\mathbf{H}_{ii}^{rr}\mathbf{G}_i^r {\mathbf{G}_i^r}^H {\mathbf{H}_{ii}^{rr}}^H)^T \\ &+ \sum_{j=1|j\neq i}^{N} \left(({\mathbf{G}_j^l}^T \mathbf{H}_{ji}^{lr})^H ({\mathbf{G}_j^l}^T \mathbf{H}_{ji}^{lr}) + ({\mathbf{G}_j^r}^T \mathbf{H}_{ji}^{rr})^H ({\mathbf{G}_j^r}^T \mathbf{H}_{ji}^{rr}) \right). \end{aligned} \tag{5}$$

$\mathbf{S}_i^r$ contains the *spatial signatures* (SPs) of the interference signal induced by radio i_r onto radio j_l ($\mathbf{H}_{ji}^{lr}$) and onto radio j_r ($\mathbf{H}_{ji}^{rr}$). It also captures the SPs of the information signals ($\mathbf{H}_{jj}^{lr}$, $\mathbf{H}_{jj}^{rl}$) intended for radios j_l and j_r that are implicitly embedded in transmit precoders $\mathbf{G}_j^l$ and $\mathbf{G}_j^r$ (as radio j_l aligns its data streams with the sub-channels directions of $\mathbf{H}_{jj}^{rl}$ while $\mathbf{H}_{jj}^{lr} = (\mathbf{H}_{jj}^{rl})^T$).

Intuitively, for an interfering channel, SPs capture the *vulnerable* directions that interference is most harmful. For an information signal, SPs are directions along which the transmit/receive beamformers should align the signal to maximize the signal's received power [19][1]. Exploiting knowledge of other nodes' SPs, learned while transmitting, an FD radio can meet its rate demand while minimizing both transmit power and interference induced on other radios. To that end, the precoder of radio i_r can be obtained by solving:

$$\begin{aligned} &\underset{\{\mathbf{G}_i^r\}}{\text{minimize}} \;\; \text{tr}(\mathbf{G}_i^r \mathbf{G}_i^{rH}) + \text{tr}(\mathbf{G}_i^r \mathbf{S}_i^r \mathbf{G}_i^{rH}) \\ &\text{s.t.} \qquad d_i^l \leq c_i^l \end{aligned} \tag{6}$$

where $\text{tr}(\mathbf{G}_i^r \mathbf{G}_i^{rH}) + \text{tr}(\mathbf{G}_i^r \mathbf{S}_i^r \mathbf{G}_i^{rH}) = \text{tr}(\mathbf{G}_i^r {\mathbf{Q}_i^r}^T \mathbf{G}_i^{rH})$ is interpreted as the summation of transmit power and interference caused by i_r.

Note that we area considering M bi-directional FD links. The key difference between this model and the case with $2M$ directional links is the channel reciprocity (thanks to the co-location of Tx and Rx of a FD radio) and the self-interference. If one was to replace our model of M bidirectional links with $2M$ directional links then the channel reciprocity does not hold. That makes the spatial signature exploitation not possible. As such, recruiting the game (6) does not lead to an energy-efficient NE as observed in our case. Additionally, using the bi-directional link model, the $\mathbf{Q}_i^l$ and $\mathbf{Q}_i^r$ (noise-plus-interference covariance matrix at radio i_l and i_r) are *readily available* at the Rx antennas without requiring any feedback from other radios in the network. We do not require any CSI feedback from neighboring links. The observation of channel reciprocity in FD radios, interference interpretation, and the game (6) follow a similar setting but of conventional HD MIMO radios in [29].

[1]For example, the minimum mean square error (MMSE) receiver at j_l is capacity-optimal when its receive beamformer is set to $[\mathbf{Q}_j^l]^{-1}\mathbf{H}_{jj}^{lr}$ [19].

B. Nash Equilibrium Existence and Uniqueness

The two games (4) and (6) have identical strategic space, defined as the union of all players' strategic spaces [30] and shaped by the rate constraints C1 and C2. We can thus focus on analyzing game (6); game (4) then follows by replacing ${\mathbf{Q}_i^r}^T$ in (6) with the identity matrix $\mathbf{I}$.

Optimizing the precoder $\mathbf{G}_i^r$ of radio i_r embodies computing the optimal radiation directions and power allocation across i_r's antennas. We can rewrite $\mathbf{G}_i^r$ as:

$$\mathbf{G}_i^r = \tilde{\mathbf{G}}_i^r \times {\mathbf{P}_i^r}^{1/2} \tag{11}$$

where $\tilde{\mathbf{G}}_i^r$ is an $M \times M$ unit-norm column matrix that controls the radiation directions of radio i_r. $\mathbf{P}_i^r$ is an $M \times M$ diagonal matrix whose diagonal element $\mathbf{P}_i^r(m,m)$ is the power allocated on the mth data stream of radio i_r.

Let $\boldsymbol{p}_i^r \overset{\text{def}}{=} [\mathbf{P}_i^r(1,1), \mathbf{P}_i^r(2,2), \ldots, \mathbf{P}_i^r(M,M)]$ denote the power allocation of radio i_r for its various data streams. Let $\boldsymbol{p} \overset{\text{def}}{=} [\boldsymbol{p}_1^l, \boldsymbol{p}_1^r, , \ldots, \boldsymbol{p}_N^l, \boldsymbol{p}_N^r] \in \mathbb{R}_+^{2NM}$ denote the power allocation on all data streams of all $2N$ radios in the network. $\mathbf{Q}_i^r$ is positive definite, and so is its transpose. The objective function in (6) is non-decreasing in every element of $\boldsymbol{p}_i^r$. Thus, at a NE of the game (if one exists), the rate demand inequality constraint becomes an equality. Otherwise, a radio would be able to lower its transmit power and reduce the objective function in (6) while fulfilling its rate demand. This defines a feasible set for $\boldsymbol{p}$, denoted by $\mathbb{P}_{feasible}(\boldsymbol{d})$ in (7), corresponding to a given requested rate profile $\boldsymbol{d} \overset{\text{def}}{=} [d_1^r, d_1^l, \ldots, d_N^r, d_N^l]$ at a NE.

Theorem 1: If a given set of rate demands (at all links) can be supported with a finite power vector $\boldsymbol{p}$, then the game (6) admits at least one NE for this set of rate demands.

Proof: A player's strategic space in Theorem 1 is defined by all possible precoding matrices $\mathbf{G}_i^r$ that satisfy the rate demand constraint $d_i^l \leq c_i^l$ of the player$i$. This strategic space is also the feasible region of the player's optimization problem (6). It is non-empty if the feasible region is non-empty or the problem (6) is feasible.

Note that in the standard game theory, it is often implicitly understood that there is always an action/strategy for a player, i.e., the strategic space is non-empty. However, for the underlying power minimization problem subject to rate demands. If the rate demands are too high, then a feasible solution may not exist.We refer to this scenario as having an empty strategic space. By contrast, for all other cases, the strategic space is said non-empty.

The NE existence proof in [30] states that a game with a concave utility function and compact and convex strategic space for each player admits at least one NE. As stipulated in Theorem 1, a given set of rate demand can be fulfilled. Hence then all per-user optimization problems in (6) are feasible, i.e.,

$$\mathbb{P}_{feasible}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{p} \in \mathbb{R}_+^{2NM} | c_i^l(\boldsymbol{p}) \stackrel{\text{def}}{=} \log |\mathbf{I} + \mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH}{\mathbf{Q}_i^l}^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r| = d_i^l, \forall i_r, \quad , i_l\right\} \tag{7}$$

$$\boldsymbol{\Gamma} \stackrel{\text{def}}{=} \begin{bmatrix} |\mathbf{H}_{11}^{lrH}\mathbf{H}_{11}^{lr}|^{\frac{1}{M}} & -g_{sis}(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{11}^{llH}\mathbf{H}_{11}^{ll})}{M} & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{12}^{llH}\mathbf{H}_{12}^{ll})}{M} & \cdots & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{llH}\mathbf{H}_{1N}^{ll})}{M} & -(2^{d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{lrH}\mathbf{H}_{1N}^{lr})}{M} \\ -g_{sis}(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{11}^{rrH}\mathbf{H}_{11}^{rr})}{M} & |\mathbf{H}_{11}^{rlH}\mathbf{H}_{11}^{rl}|^{\frac{1}{M}} & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{12}^{rlH}\mathbf{H}_{12}^{rl})}{M} & \cdots & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{rlH}\mathbf{H}_{1N}^{rl})}{M} & -(2^{d_1^r}-1)\frac{\text{tr}(\mathbf{H}_{1N}^{rrH}\mathbf{H}_{1N}^{rr})}{M} \\ \vdots & \vdots & \ddots & \vdots & & \\ -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1}^{rlH}\mathbf{H}_{N1}^{rl})}{M} & -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1}^{rrH}\mathbf{H}_{N1}^{rr})}{M} & -(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N2}^{rlH}\mathbf{H}_{N2}^{rl})}{M} & \cdots & -g_{sis}(2^{d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{NN}^{rrH}\mathbf{H}_{NN}^{rr})}{M} & |\mathbf{H}_{NN}^{rlH}\mathbf{H}_{NN}^{rl}|^{\frac{1}{M}} \end{bmatrix} \tag{8}$$

$$\mathbb{P}_{asymp}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{f} \in \mathbb{R}_+^{2NM} | \exists\{\boldsymbol{p}_n\} \in \mathbb{P}_{feasible}(\boldsymbol{d}) \text{ and } \{\nu_n\} \to +\infty \text{ so that } \lim_{n\to\infty}\frac{\boldsymbol{p}_n}{\nu_n} = \boldsymbol{f}\right\} \tag{9}$$

$$\mathbb{P}(\boldsymbol{d}) \stackrel{\text{def}}{=} \left\{\boldsymbol{f} \in \mathbb{R}_+^{2NM} | {c'}_i^l(\boldsymbol{f}) \stackrel{\text{def}}{=} \log\left(1 + \frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l) + \sum\limits_{j=1|j\neq i}^{N}\left(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l) + \frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r)\right)}\right) \leq d_i^l, \forall i\right\} \tag{10}$$

their strategic spaces are non-empty. Additionally, the feasible region (or the strategic space) of each player in problem (6) is convex, as the achievable throughput c_i^l for each player i is concave w.r.t. the radio's precoder $\mathbf{G}_i^r$ [19]. The only remaining requirement is the compactness of the strategic space. The precoder (or strategy) $\mathbf{G}_i^r$, rewritten as $\mathbf{G}_i^r = \tilde{\mathbf{G}}_i^r \times {\mathbf{P}_i^r}^{1/2}$ is finite/bounded because the elements in the power allocation matrix $\mathbf{P}_i^r$ are bounded. As a set of rate demands can be met with finite power, these technical constraints can be set by a sufficiently high, yet finite, power. In short, the strategic space of problem (6) is nonempty, convex, and compact. Moreover, the player in problem (6) minimizes a convex objective function (by verifying its Hessian is positive semidefinite w.r.t. $\mathbf{G}_i^r$), i.e., the player maximizes a concave utility function. Citing [30], problem (6) admits at least one NE. □

For the existence of a NE, it suffices to find conditions under which the feasible set $\mathbb{P}_{feasible}(\boldsymbol{d})$ of $\boldsymbol{p}$ is nonempty and bounded. This is formally stated in the following theorem:

Theorem 2: Let $\boldsymbol{\Gamma}$ be the $2N \times 2N$ matrix defined in (8). If $\boldsymbol{\Gamma}$ is a P-matrix[2], then $\mathbb{P}_{feasible}(\boldsymbol{d}) \in \mathbb{R}_+^{2NM}$ is nonempty and bounded, and hence game (6) admits at least one NE.

Proof: We first prove that $\mathbb{P}_{feasible}(\boldsymbol{d})$ contains at least one bounded vector $\boldsymbol{p} \in \mathbb{R}_+^{NKM}$ or the rate remand can be met with finite transmit power.

Lemma 1: Given that $\boldsymbol{\Gamma}$ is a P-matrix, there exists at least one bounded vector $\boldsymbol{p} \in \mathbb{P}_{feasible}(\boldsymbol{d}) \in \mathbb{R}_+^{2NM}$.

Proof: See Appendix A. □

Next, to show that $\mathbb{P}_{feasible}(\boldsymbol{d})$ is bounded, we rely on the concept of asymptotic cone of a nonempty set in recession analysis [27]. Specifically, for a nonempty set $\mathbb{P} \in \mathbb{R}_+^N$, its *asymptotic cone*, $\mathbb{P}_{asymp}$, is comprised of vectors $\boldsymbol{f} \in \mathbb{R}_+^N$, called limit directions. Each limit direction vector $\boldsymbol{f}$ is defined through the existence of a sequence of vectors $\boldsymbol{p}_n \in \mathbb{P}$ and a sequence of scalars ν_n that tend $+\infty$ such that [27]:

$$\lim_{n\to\infty}\frac{\boldsymbol{p}_n}{\nu_n} = \boldsymbol{f}. \tag{12}$$

$\mathbb{P}$ is bounded if its asymptotic cone $\mathbb{P}_{asymp} = \{\mathbf{0}\}$ [27]. To show that $\mathbb{P}_{feasible}(\boldsymbol{d})$ is bounded, it suffices to prove that its asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$ contains only the zero vector. The asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$ is formally defined in (9).

Since $\mathbb{P}_{feasible}(\boldsymbol{d})$ has at least one bounded $\boldsymbol{p}$ (Lemma 1), by the definition of limit directions, the vector zero $\mathbf{0}$ belongs to its asymptotic cone $\mathbb{P}_{asymp}(\boldsymbol{d})$. We now construct a set $\mathbb{P}(\boldsymbol{d})$ of which $\mathbb{P}_{asymp}(\boldsymbol{d})$ is a subset and prove that $\mathbb{P}(\boldsymbol{d}) = \{\mathbf{0}\}$ if $\boldsymbol{\Gamma}$ is a P-matrix.

Lemma 2: If $\boldsymbol{f} \in \mathbb{P}_{asymp}(\boldsymbol{d})$ then $\boldsymbol{f}$ belongs to $\mathbb{P}(\boldsymbol{d})$, defined in (10).

Proof: See Appendix B. □

Assuming that there exists at least one $\boldsymbol{f} \neq \mathbf{0}$ and that $\boldsymbol{f} \in \mathbb{P}(\boldsymbol{d})$, we have:

$$\boldsymbol{\Gamma} \times [\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l), \ldots, \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T \leq \mathbf{0}. \tag{13}$$

As $\boldsymbol{\Gamma}$ is a P-matrix and $[\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l), \ldots, \text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T$ is a nonnegative vector, (13) implies that $\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l) = 0$ and $\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r) = 0$ $\forall i$ [31] or $\boldsymbol{f} = \mathbf{0}$. This contradicts the above assumption. Hence, $\mathbb{P}(\boldsymbol{d})$ and its subset $\mathbb{P}_{asymp}(\boldsymbol{d})$ equal to $\{\mathbf{0}\}$. Theorem 2 is proved. □

Intuitions behind Theorem 2 can be drawn as follows. If the diagonal elements of $\boldsymbol{\Gamma}$ are positive, then a sufficient condition for $\boldsymbol{\Gamma}$ to be a P-matrix is $|\boldsymbol{\Gamma}(i,i)| \geq \sum\limits_{j\neq i}|\boldsymbol{\Gamma}(i,j)|$ (i.e., row diagonally dominant) [31]. Hence, the following inequality guarantees that game (6) has at least one NE:

$$\frac{M\det(\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr})^{\frac{1}{M}}}{g_{sis}\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll}) + \sum\limits_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}) + \text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}))} \geq (2^{d_i^l}-1), \forall i \in \mathcal{N}. \tag{14}$$

To better interpret inequality (14), let's assume perfect SIS (i.e., g_{sis} is sufficiently small to be neglected) and rewrite $\mathbf{H}_{ii}^{lr} = \frac{1}{\sqrt{s_{ii}^{lr\,n}}}\bar{\mathbf{H}}_{ii}^{lr}$, where n is the path loss exponent, ${s_{ii}}^{lr}$ is the transmission distance from radio i_r to radio i_l, and $\bar{\mathbf{H}}_{ii}^{lr}$ is a complex Gaussian matrix with with zero mean and unit

[2]A P-matrix is one for which all principal minors are positive [31].

variance. Inequality (14) can be rewritten as:

$$\frac{M\det(\bar{\mathbf{H}}_{ii}^{lrH}\bar{\mathbf{H}}_{ii}^{lr})^{\frac{1}{M}}}{\sum\limits_{j=1|j\neq i}^{N}((\frac{s_{ii}^{lr}}{s_{ij}^{ll\,n}})^{n}\mathrm{tr}(\bar{\mathbf{H}}_{ij}^{llH}\bar{\mathbf{H}}_{ij}^{ll})+(\frac{s_{ii}^{lr}}{s_{ij}^{lr}})^{n}\mathrm{tr}(\bar{\mathbf{H}}_{ij}^{lrH}\bar{\mathbf{H}}_{ij}^{lr}))}\geq(2^{{d_i}^{l}}-1)\forall\ i. \tag{15}$$

The nominator of the LHS in (15) represents the strength of the channel gain matrix from i_r to i_l, while its denominator describes the strength of (interfering) channel gain matrices from all other radios j_l and j_r ($j \neq i$) on radio i_l. For the game in (6) to have at least one NE, the multi-user interference should not be too strong. This is the case if the (transmission) distance s_{ii}^{lr} between i_l and i_r is small enough compared with (interfering) distances (s_{ij}^{ll} and s_{ij}^{lr}) between i_l and radios other than i_r, the channel gain matrix of link i is full-rank (this is often the case in a rich-scattering environment), and its requested rate is not too high. The acceptable multi-user interference is explicitly quantified in (14), and is a function of the rate demand d_i^l of radio i_l. For higher rate demands, inequality (14) becomes stringent, meaning that network interference must be lowered.

Remark 1: The denominator of (15) captures the interference to i_l from radios on the left side j_l and right side j_r sides of all links $j \neq i$. If all FD radios choose to operate in HD fashion, e.g., all left radios transmit and all right radios receive, then the second term in the denominator should disappear. Consequently, the denominator reduces roughly by half. If (15) holds for the new denominator for all radios with rate demands $d_i^r = d$ and $d_i^l = 0$ (i.e., all left radios are transmitters only) for all i, then when all radios return to FD mode, (15) should also hold for all radios with rate demands $d_i^r = d-1$ and $d_i^r = d-1$. This is because $\frac{(2^d-1)}{2} > 2^{d-1}-1$. The network throughput in the FD case is then $2N(d-1)$, which is asymptotically twice that of the HD case (Nd). Hence, we can conjecture that if (15) holds for all radios, the network throughput can double with FD radios. Note that, in general, $d \gg 1$ ($d = 1$ corresponds to the extreme case when the signal-to-noise-plus-interference ratio is 1 or signal strength is equal to the noise plus interference). Conditions in (15) are also in line with the findings in [9] [10], where it was observed that FD radios outperform HD ones if network interference is mild (i.e., interfering links are sufficiently separated from each other). However, [9] [10] did not quantify how mild network interference should be for FD radios to double the network throughput.

To analyze the uniqueness of the NE, we rely on variational inequalities (VI) theory, casting (6) as a VI problem. A tutorial on VI can be found in [28] and the references therein.

Theorem 3: If game (6) has a NE, this NE is unique.

Proof: We prove that the mapping of the equivalent VI problem of (6) is continuous uniformly-P function. Hence, if a NE exists, it is unique. See Appendix C for details. □

Remark 2: The conditions in Theorem 2 or inequality (15) are sufficient for throughput doubling but not necessary. FD radios can still double network throughput even when these conditions do not hold. Theorem 3 indicates that (6) does not have multiple NEs. In our simulations, we observed that the game always converges to its unique NE as long as the rate demands are not unreasonably high.

C. Best Response

The optimal precoder $\mathbf{G}_i^r$ of radio i_r is obtained by solving (6). Notice that (6) is convex, hence can be solved efficiently using the interior-point method. To gain insight into how power is allocated over i_r's antennas, we follow the approach in [29]. Specifically, the Lagrange function of (6) is:

$$\begin{aligned}
&L_i^r(\mathbf{G}_i^r,\gamma_i^l)\\
&=\mathrm{tr}(\mathbf{G}_i^r\mathbf{Q}_i^{rT}\mathbf{G}_i^{rH})+\gamma_i^l(d_i^l-\log|\mathbf{I}+\mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r|)\\
&=\mathrm{tr}(\bar{\mathbf{G}}_i^r\bar{\mathbf{G}}_i^{rH})+\gamma_i^l(d_i^l-\log|\mathbf{I}+\bar{\mathbf{G}}_i^{rH}\mathbf{E}_i^{r-1}\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}\mathbf{E}_i^{rH^{-1}}\bar{\mathbf{G}}_i^r|)\\
&\leq\mathrm{tr}(\bar{\mathbf{G}}_i^r\bar{\mathbf{G}}_i^{rH})\\
&\quad+\gamma_i^l(d_i^l-\sum_{m=1}^{M}\log(1+\mathrm{diag}_m(\bar{\mathbf{G}}_i^{rH}\mathbf{E}_i^{r-1}\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}\mathbf{E}_i^{rH^{-1}}\bar{\mathbf{G}}_i^r|)))
\end{aligned} \tag{16}$$

where γ_i^l and γ_i^r are nonnegative Lagrangian multipliers, and $\bar{\mathbf{G}}_i^{rH} = \mathbf{G}_i^{rH}\mathbf{E}_i^r$ with Cholesky decomposition $\mathbf{Q}_i^{rT} = \mathbf{E}_i^r\mathbf{E}_i^{rH}$, and $\mathrm{diag}_m(*)$ denotes the diagonal element (m, m) of the matrix. The last inequality is obtained by applying Hadamard inequality [32].

Problem (6) can be solved by finding the maximum of its lower bound $L_i^r(\tilde{\mathbf{G}}_i^r, \gamma_i^l)$. Inequality (16) becomes an equality if there exists an orthonormal matrix $\bar{\mathbf{G}}_i^r$ that diagonalizes $\mathbf{E}_i^{r-1}\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}\mathbf{E}_i^{rH^{-1}}$. After some manipulations, we can prove that the optimal $\mathbf{G}_i^r$ must be in the form a generalized eigen matrix of $\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}$ and $\mathbf{Q}_i^{rT}$. This is realized by setting $\tilde{\mathbf{G}}_i^r$ in (11) as a unit-norm generalized eigen matrix of $\mathbf{H}_{ii}^{lrH}[\mathbf{Q}_i^l]^{-1}\mathbf{H}_{ii}^{lr}$ and $\mathbf{Q}_i^{rT}$. It follows from [33] that:

$$\begin{aligned}
&\tilde{\mathbf{G}}_i^{rH}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}\tilde{\mathbf{G}}_i^r = \mathbf{\Pi}_i^l \text{ and}\\
&\tilde{\mathbf{G}}_i^{rH}\mathbf{Q}_i^{rT}\tilde{\mathbf{G}}_i^r = \mathbf{\Omega}_i^r
\end{aligned} \tag{17}$$

where $\mathbf{\Pi}_i^l$ and $\mathbf{\Omega}_i^r$ are diagonal matrices.

The Lagrangian $L_i^r(\mathbf{G}_i^r, \gamma_i^l)$ becomes:

$$\begin{aligned}
&L_i^r(\mathbf{G}_i^r,\gamma_i^l)\\
&=\sum_{m=1}^{M}(\mathrm{diag}_m(\mathbf{\Omega}_i^r)\mathbf{P}_i^r(m,m)-\gamma_i^l\log(1+\mathbf{P}_i^r(m,m)\mathrm{diag}_m(\mathbf{\Pi}_i^l))).
\end{aligned}$$

The optimal power for data stream m is obtained by equating the derivative of $L_i^r(\mathbf{G}_i^r, \gamma_i^l)$ to zero. Accordingly:

$$\mathbf{P}_i^r(m,m) = \max\left(0, \gamma_i^l\frac{1}{\mathrm{diag}_m(\mathbf{\Omega}_i^r)} - \frac{1}{\mathrm{diag}_m(\mathbf{\Pi}_i^l)}\right) \tag{18}$$

where the Lagrange multiplier γ_i^l is computed (e.g., using bisection search) to meet the rate demand d_i^l.

From (18), more power is allocated on data streams with lower $\mathrm{diag}_m(\mathbf{\Omega}_i^r)$ and higher $\mathrm{diag}_m(\mathbf{\Pi}_i^l)$. This means more power is allocated to higher-gain streams and less power on directions that cause higher interference to others.

$\mathbf{Q}_i^l$ and $\mathbf{Q}_i^r$ in Equations (17) and (18) are the noise-plus-interference covariance matrix at radio i_l and i_r. These quantities are *readily available* at the Rx antennas without requiring any feedback from other radios in the network. To compute the precoder at a radio of a link i, no CSI between the radio of link i and any other radio of link j, $j \neq i$, is required. One only needs the local CSI of link i that can be easily

$$\mathbf{\Gamma}_k \stackrel{\text{def}}{=} \begin{bmatrix} |\mathbf{H}_{11k}^{lrH}\mathbf{H}_{11k}^{lr}|^{\frac{1}{M}} & -g_{sis}(2^{\lambda_{1k}^l d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{11k}^{llH}\mathbf{H}_{11k}^{ll})}{M} & \cdots & \cdots & -(2^{\lambda_{1k}^l d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1Nk}^{llH}\mathbf{H}_{1Nk}^{ll})}{M} & -(2^{\lambda_{1k}^l d_1^l}-1)\frac{\text{tr}(\mathbf{H}_{1Nk}^{lrH}\mathbf{H}_{1Nk}^{lr})}{M} \\ -g_{sis}(2^{\lambda_{2k}^r d_2^r}-1)\frac{\text{tr}(\mathbf{H}_{11k}^{rrH}\mathbf{H}_{11k}^{rr})}{M} & |\mathbf{H}_{11k}^{rlH}\mathbf{H}_{11k}^{rl}|^{\frac{1}{M}} & \cdots & \cdots & -(2^{\lambda_{2k}^r d_2^r}-1)\frac{\text{tr}(\mathbf{H}_{1Nk}^{rlH}\mathbf{H}_{1Nk}^{rl})}{M} & -(2^{\lambda_{2k}^r d_2^r}-1)\frac{\text{tr}(\mathbf{H}_{1Nk}^{rrH}\mathbf{H}_{1Nk}^{rr})}{M} \\ \vdots & \vdots & \ddots & \vdots & & \\ -(2^{\lambda_{Nk}^r d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1k}^{rlH}\mathbf{H}_{N1k}^{rl})}{M} & -(2^{\lambda_{Nk}^r d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{N1k}^{rrH}\mathbf{H}_{N1k}^{rr})}{M} & \cdots & \cdots & -g_{sis}(2^{\lambda_{Nk}^r d_N^r}-1)\frac{\text{tr}(\mathbf{H}_{NNk}^{rrH}\mathbf{H}_{NNk}^{rr})}{M} & |\mathbf{H}_{NNk}^{rlH}\mathbf{H}_{NNk}^{rl}|^{\frac{1}{M}} \end{bmatrix} \quad (20)$$

acquired using various channel estimation and training methods (similar to that in existing systems, e.g., 802.11ac, LTE). MIMO CSI estimation and training have a well-established literature, hence we choose not to dwell on it. Note that estimating the CSI for the "main channel" between a Tx and its intended Rx" is also needed for HD MIMO networks.

D. Multi-carrier FD MIMO

In this section, we extend the above setup and results to multi-carrier setting (e.g., OFDM with channel bonding/aggregation). Assume that on each direction, a FD radio can transmit simultaneously over K channels/subcarriers in Φ_K. Let $\mathbf{H}_{iik}^{rl}$ ($\mathbf{H}_{iik}^{lr}$) denote the $M \times M$ channel gain matrix of the left-to-right (right-to-left) direction of link i and $\mathbf{G}_{ik}^{l}$ ($\mathbf{G}_{ik}^{r}$) denote the transmit precoding matrices at the left i_l (right i_r) radios of link i on channel $k \in \Phi_K$. The game in (6) can be rewritten as follows for the case of K carriers:

$$\begin{aligned} & \underset{\{\mathbf{G}_{ik}^r\}}{\text{minimize}} \sum_{k=1}^{K} \text{tr}(\mathbf{G}_{ik}^r \mathbf{G}_{ik}^{rH}) + \text{tr}(\mathbf{G}_{ik}^r \mathbf{S}_{ik}^r \mathbf{G}_{ik}^{rH}) \\ & \text{s.t.} \qquad d_i^l \le c_{iK}^l \end{aligned} \quad (19)$$

where $\mathbf{S}_{ik}^r$ is equivalent to $\mathbf{S}_i^r$ in (5) but over channel k and $c_{iK}^l \stackrel{\text{def}}{=} \sum_{k=1}^{K} c_{ik}^l$ with c_{ik}^l is the achieved rate of i_l on channel k.

The following theorem states the sufficient conditions for the existence of a unique NE.

Theorem 4: If there exists a non-negative $1 \times 2NK$ vector $\lambda \stackrel{\text{def}}{=} \{\lambda_{11}^l, \lambda_{11}^r, \ldots, \lambda_{1k}^l, \lambda_{1k}^r, \ldots, \lambda_{NK}^l, \lambda_{NK}^r\}$ such that $\sum_{k=1}^{K} \lambda_{ik}^r = 1, \sum_{k=1}^{K} \lambda_{ik}^l = 1, \forall i = 1...N$ and $\mathbf{\Gamma}_k$ (defined in (20)) is a P-matrix for all $k = 1...K$, then game (19) admits a unique NE.

Proof:

First, consider a channel k. As $\mathbf{\Gamma}_k$ is a P-matrix, Theorem 2 implies that rates $\{\lambda_{1k}^l d_1^l, \lambda_{1k}^r d_1^r, \ldots, \lambda_{Nk}^l d_k^l, \lambda_{Nk}^r d_k^r\}$ are achievable on channel k. Thus, for $\sum_{k=1}^{K} \lambda_{ik}^r = 1, \sum_{k=1}^{K} \lambda_{ik}^l = 1, \forall i = 1...N$ and that $\mathbf{\Gamma}_k$ is a P-matrix for all $k = 1...K$, the rates $[d_1^r, d_1^l, \ldots, d_N^r, d_N^l]$ are achievable by aggregating over all K channels under the game (19). Analogous to Theorem 1, we can easily prove (by appealing the convex/concave game analysis) that if the rates $[d_1^r, d_1^l, \ldots, d_N^r, d_N^l]$ can be supported, then game (19) admits at least one NE. To prove that the NE is in fact unique, one can follow the routine of Theorem 3's proof using VI theory. □

E. MAC Protocol

We briefly present a MAC protocol, called FD-MAC, that implements game (6) in a distributed fashion. Unlike typical CSMA-based protocols, FD-MAC exploits information perceived by FD radios to enable concurrent transmissions on multiple links. Each transmission session in FD-MAC consists of two phases: a training phase and a data phase. In the first phase, an FD radio, say A, that has data packets to send, starts by transmitting a hand-shaking message (HSK). This HSK contains a training sequence for CSI estimation purposes (to rendezvous with its intended radio B). HSKs are sent at the lowest (most robust) rate, referred to as a base rate, so as to improve the chances of delivering them.

As each FD radio can transmit and receive at the same time and a two-way channel exists between the two radios of a bidirectional link, FD radios of a link can instantaneously update each other regarding CSI as well as the noise-plus-interference covariance matrix. This information is needed to solve (6). If either A or B fails to transmit or receive at the base rate to hand-shake with the intended partner, it then backs off for a random duration before trying again. Upon receiving an HSK, radio B replies with a message to trigger the training process by solving problem (6) to achieve the rate demand. The data transmission phase ensues with multiple back-to-back packets. Note that under FD-MAC, radios of different links do not need to coordinate or exchange any signaling packets but precoders are computed on the fly using *only local information*.

To ensure that the training phase in FD-MAC is not too long, the iterating process should converge after a reasonable time. The iterative process of updating the best responses of players/links can be done in a synchronous or asynchronous manner. In the former, players update and execute their best responses/strategies either in a sequential (i.e., Gauss-Seidel) or parallel (i.e., Jacobi) manner[3]. To facilitate synchronous update, players have to coordinate with each other or must be in sync. Such a requirement is quite challenging due to the dynamic nature of the network topology. The Jacobi, Gauss-Seidel or synchronous updates are special forms of the asynchronous update scheme in which players are allowed to update their best responses in an arbitrary manner, or even sporadically skip their responses. To prove the convergence of the best response in FD-MAC to the unique NE under the asynchronous updates, one can rely on the Asynchronous Convergence Theorem in [34]. For brevity, we omit this detailed proof. Interested readers are referred to a similar proof in [35] and therein references.

IV. NETWORK-WIDE PROBLEM

Note that even if the network-wide problem can be optimally solved in a centralized manner, collecting the global network information for this network-wide problem requires excessive amount of overhead. As such, finding a distributed and optimal centralized solution to this network-wide problem is, yet very much desirable, remains an open problem.

[3]The parallel (or Jacobi) update is often observed to converge faster than the sequential (or Gauss-Seidel) one [28].

To seek a performance benchmark for our distributed solution above, in this section we use the augmented Lagrange multiplier method [36] to derive the centralized algorithm for the network-wide problem (3). The centralized algorithm provides locally optimal solution. The augmented Lagrange of (3) is given in (23), where $q_i^l \stackrel{\text{def}}{=} d_i^l - c_i^l$, $q_i^r \stackrel{\text{def}}{=} d_i^r - c_i^r$ and p is a positive penalty factor for violating rate constraints. At an optimal solution, (24) holds.

Since q_j^l is continuously differentiable w.r.t every entry of $\tilde{\mathbf{G}}_j^r$, the third and fourth terms in (24) are also continuously differentiable [36]. Their derivatives are as follows:

$$\frac{\partial\{(\max\{0,\gamma_i^l+pq_i^l\})^2\}}{\partial \mathbf{G}_j^{r*}}\}=\begin{cases}0 & \text{if } \gamma_i^l+pq_i^l \le 0\\ -2p\frac{\partial c_i^l}{\partial \mathbf{G}_i^{r*}} & \text{if } \gamma_i^l+pq_i^l > 0 \text{ and } i=j\\ -2p\frac{\partial c_i^l}{\partial \mathbf{G}_j^{r*}} & \text{if } \gamma_i^l+pq_i^l > 0 \text{ and } i\neq j\end{cases}$$

$$\frac{\partial\{(\max\{0,\gamma_i^r+pq_i^r\})^2\}}{\partial \mathbf{G}_j^{r*}}\}=\begin{cases}0 & \text{if } \gamma_i^r+pq_i^r \le 0 \text{ or } j=i\\ -2p\frac{\partial c_i^r}{\partial \tilde{\mathbf{G}}_j^{r*}} & \text{if } \gamma_i^r+pq_i^r > 0 \text{ and } j\neq i\end{cases}$$

where:

$$\frac{\partial c_i^l(\mathbf{G})}{\partial \mathbf{G}_i^{r*}} = -\mathbf{H}_{ii}^{lrH}(\mathbf{Q}_i^l + \mathbf{H}_{ii}^{lr}\mathbf{G}_i^r\mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH})^{-1}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r \quad (21)$$

and

$$\begin{aligned}\frac{\partial c_i^l(\mathbf{G})}{\partial \mathbf{G}_j^{r*}} =& \mathbf{H}_{ij}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ij}^{lr}[(\mathbf{G}_i^r\mathbf{G}_i^{rH})^{-1}\\ &+ \mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}]^{-1}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^r\end{aligned} \quad (22)$$

We use the gradient search algorithm with Armijo step size [36] to find $(\mathbf{G}_i^l, \mathbf{G}_i^r, \gamma_i^l, \gamma_i^r, p)$ such that (24) holds for all radios. The running time can be high as it involves NM^2 complex variables (or $2NM^2$ real ones).

V. Simulations Results

We numerically evaluate the performance of the above centralized and distributed algorithms using MATLAB simulations. We compare these algorithms in terms of the total network power required to meet a given set of rate demands when the FD capability to capture spatial signatures is exploited (game (6), FD-MAC) and when it is not exploited (game (4), referred to as "without SPE"). We also compare the total required power under the proposed FD-MAC protocol with that when FD links take turn to access the channel (e.g., CSMA-based protocols). Eight pairs of radios are randomly placed in a field of dimensions 1000×1000 m^2. Each radio has 4 antennas. Channel bandwidth is 20 MHz. We assume perfect SIS with $g_{sis} = 0$. Noise floor is set as -90dBm/Hz. The channel fading is flat with a free-space attenuation factor of 2. All algorithms have identical initializations of precoding matrices.

Fig. 2 shows snapshots of radiation patterns of FD radios under different algorithms (at their converged points). The digital beamforming radiation pattern at each node is a function of the node's precoder, e.g., matrix $\tilde{\mathbf{G}}_i^r$. We first normalize the matrix with its norm and then compute the distance ρ in the polar coordinate from this normalized precoder. After that, we convert the polar coordinates to Cartesian coordinates to plot the pattern. As can be seen, when nodes cooperate so that (3) can be solved in a centralized manner, we visually notice that FD radios try to steer their beams away from unintended receivers. This is also observed for the distributed FD-MAC algorithm with SPE, where nodes minimize their Tx power weighted by SPs of others. Interestingly, the beam patterns of the distributed FD-MAC are quite similar to that of the centralized algorithm, suggesting the efficiency of using SPs. These two algorithms seem to induce less network interference, compared with the beam patterns when SPs are not exploited (game (4)).

In the following power comparison, we choose to focus on the transmit power at the air interface that is much more significant than the circuitry energy, especially for medium and long distances. It is also worth noting that most existing FD transceivers, e.g., [6] [7] [5], share similar RF components (e.g., frequency synthesizer and the power amplifier) with HD ones. These account for more than 85% of the circuitry energy [37]. Moreover, the extra power for the self-interference suppression/canceller of FD radios is quite negligible [38].

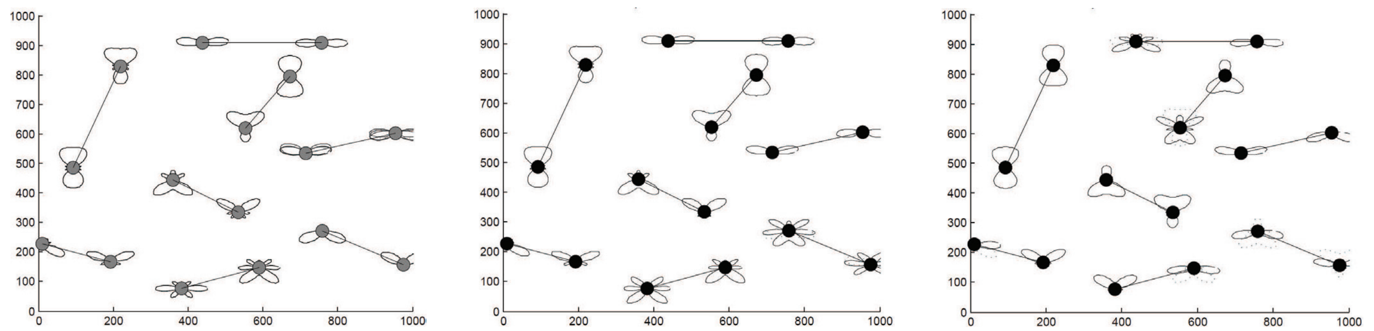
Fig. 3 depicts the total network transmit power when nodes demand a rate of 40 Mbps (i.e., 2 bps/Hz). Notice that the CSMA-based approach where FD links take turn in accessing the channel (i.e., only one FD link operates at a time), requires the least Tx power. This is because in such cases, a link does not need to cope with interference from others, but obviously this occurs at the expense of low network throughput (equals to that of one FD link's, 80 Mbps). The proposed FD-MAC algorithm converges after about 9 iterations and consumes almost the same total power as the centralized algorithm. This is in line with the similarity in the radiation behavior observed in Fig. 2. Being compared with the CSMA-based approach, by advocating concurrent links' transmission, FD-MAC requires about 5 times higher Tx power but attains 640 Mbps network

$$L(\mathbf{G}_i^l,\mathbf{G}_i^r,\gamma_i^l,\gamma_i^r,p) = \sum_{i=1}^{N}\{\text{tr}(\mathbf{G}_i^l\mathbf{G}_i^{lH})+\text{tr}(\mathbf{G}_i^r\mathbf{G}_i^{rH})+\frac{p}{2}((\max\{0,\gamma_i^l+pq_i^l\})^2-(\gamma_i^l)^2+(\max\{0,\gamma_i^r+pq_i^r\})^2-(\gamma_i^r)^2)\} \quad (23)$$

$$0 = \frac{\partial L(\mathbf{G}_i^l,\mathbf{G}_i^r,\gamma_i^l,\gamma_i^r,p)}{\partial \mathbf{G}_j^{r*}} = 2\mathbf{G}_j^r + \frac{p}{2}\sum_{i=1}^{N}\{\frac{\partial(\max\{0,\gamma_i^l+pq_i^l\})^2}{\partial \mathbf{G}_j^{r*}} + \frac{\partial(\max\{0,\gamma_i^r+pq_i^r\})^2}{\partial \mathbf{G}_j^{r*}}\} \quad (24)$$

$$\boldsymbol{x}_i^l=\left[\Re[vec(\mathbf{G}_i^l)]^T,\Im[vec(\mathbf{G}_i^l)]^T\right]^T;\boldsymbol{x}_i^r=\left[\Re[vec(\mathbf{G}_i^r)]^T,\Im[vec(\mathbf{G}_i^r)]^T\right]^T \quad (25)$$

$$\nabla_x L = 2\left[\Re[vec(\frac{\partial L}{\partial \mathbf{G}_1^{l*}})]^T,...,\Re[vec(\frac{\partial L}{\partial \mathbf{G}_N^{r*}})]^T,\Im[vec(\frac{\partial L}{\partial \mathbf{G}_1^{l*}})]^T,...,\Im[vec(\frac{\partial L}{\partial \mathbf{G}_N^{r*}})]^T\right]^T \quad (26)$$



(a) Centralized Algorithm (b) FD-MAC with SPE (c) Without SPE

Fig. 2. A snapshot of antenna patterns of FD radios under the centralized, distributed FD-MAC (exploiting SPs), and without exploiting SPs algorithms.

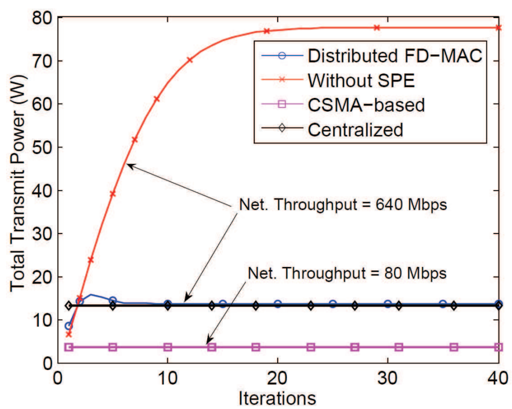


Fig. 3. Total transmit power under different algorithms and a rate demand of 40 Mbps per node.

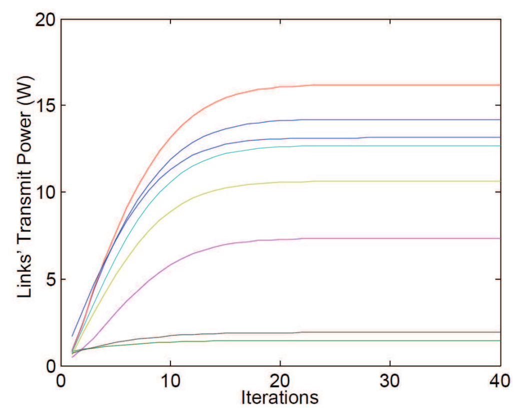


Fig. 5. Convergence of transmit power at different links when SPEs are not exploited (game (4)).

throughput (8 times higher). This gain is very significant due to the fact that throughput/rate does not scale linearly w.r.t. transmit power.

In contrast to the CSMA approach, in FD-MAC a link does not give up when the medium is busy. Instead, it proceeds but in a "responsible" way by exploiting SPs to minimize interference to ongoing receptions. We also observe that for the same amount of throughput (640 Mbps), if SPs are not exploited, the required transmit power is 77.4 W (compared to 13.1 W under FD-MAC).

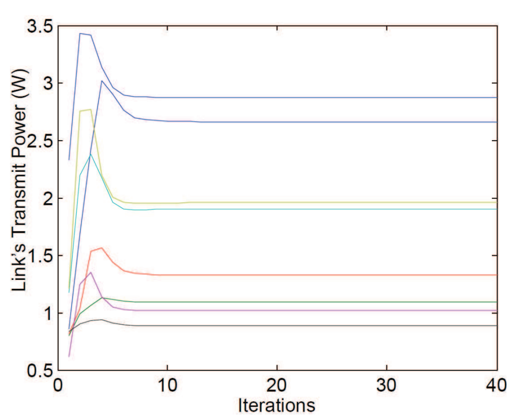


Fig. 4. Convergence of transmit power at different links under FD-MAC.

The convergence of Tx power for different links under FD-MAC is shown in Fig. 4. We notice that the consumed power values converge with different speeds (to the NE) but all intermediate Tx powers are in a similar range of the Tx power under the CSMA-based approach (Table I). This is critically important: If intermediate Tx power are excessive e.g., higher than nodes' power budget (like the case when SPs are not exploited, in Table I and Fig. 5), nodes can't follow the game to reach the NE, even if the NE itself is power-efficient. Fig. 4 and Fig. 5 (and Table I) show that by being responsible for their interference (using (6) instead of (4)) all links can reduce their Tx power. It is also seen that the nodes of game (4) (Fig. 5) take longer time to converge, compared with (6) (Fig. 4), as higher network interference makes them more dependent on each other and need more time to "negotiate". Note that the transmit power in Fig. 5 is excessively high, compared with the conventional maximum transmit power of wireless devices (about 1W). In existing radio designs, multiple radios are often not allowed to simultaneously transmit (but follow a MAC protocol). As such, the excessive transmit power in Fig. 5 would not happen in practice. However, from the simulated scenario, we observe that under our proposed scheme, FD-MAC can sustain simultaneous transmissions of FD-MIMO radios with reasonable transmit power (as in Fig. 4, always less than 3W or less than 2W on average).

To investigate the power/energy efficiency, Fig. 6 compares

TABLE I
LINKS' TX POWER (IN WATTS).

Links	HD	CSMA-based	FD-MAC	without SPE
1	1.45	0.856	2.66	14.2
2	0.0.46	0.744	1.08	1.44
3	0.39	0.29	1.32	16.2
4	0.73	0.268	1.89	12.6
5	0.24	0.327	1.01	7.3
6	0.26	0.46	1.96	10.62
7	0.322	0.43	0.88	1.9
8	0.642	0.63	2.87	13.1

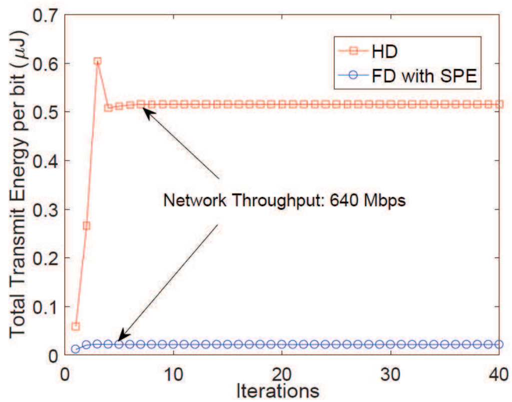


Fig. 6. Total transmit energy per bit of a HD v.s. a FD network (for the same throughput of 640 Mbps).

the total transmit power of HD (under NOMA) vs. FD radios (under FD-MAC) for the same network throughput of 640 Mbps. To that end, the rate demand for each direction of each FD link is 40 Mbps and the rate demand for each HD link is 80 Mbps. As can be seen, FD-MAC with its better capability in managing network interference and higher spectral efficiency requires much less transmit power than that of HD radios for the same network throughput. Note that the energy efficiency per bit of FD-MIMO is thanks to its spectral efficiency. However, translating this spectral efficiency at the PHY layer (of a single link) to the overall (network) energy efficiency is not trivial (as observed in Fig. 3 and its discussion) and that is the main contribution of our work (the FD-MAC protocol with the SPE).

We also observe that, when the network becomes denser, i.e., with more links (as shown in Fig. 7), FD becomes more and more energy-efficient.

To further evaluate the network interference, we adopt the network interference function (NIF) introduced in [39], as follows:

$$\text{NIF} \triangleq \text{tr}\{\sum_{i \in \mathcal{N}} (\mathbf{Q}_i^r + \mathbf{Q}_i^l)\} \tag{27}$$

Fig. 8 compares the network interference under the FD-MAC and the conventional FD (without SPE) protocols for FD radios, and HD (under NOMA) radios using the above scenario with the same network throughput of 640 Mbps. We observe FD-MAC that consumes the least power also yields the least network interference. In the case of HD, its radios require the highest transmit power (Fig. 6) but then cause the largest network interference.

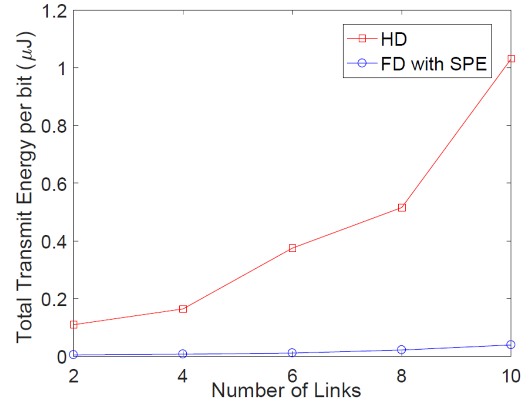


Fig. 7. Total transmit energy per bit of a HD v.s. a FD network w.r.t. the number of links (for the same throughput of 640 Mbps).

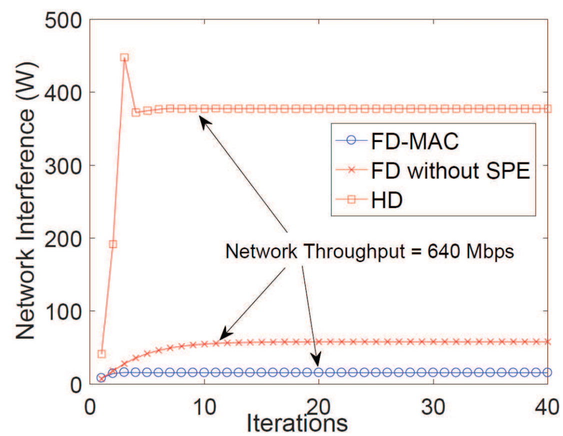


Fig. 8. Network interference comparison of a HD, a FD-MAC, and a FD-without-SPE network for the same throughput of 640 Mbps.

VI. CONCLUSIONS

To leverage the potential of FD MIMO radios at the network layer, we first investigated the transmit power minimization problem of an FD MIMO network subject to rate demands under a noncooperative game framework. We then established the conditions under which FD radios can asymptotically double the network throughput over HD ones. These conditions quantify the exact level of network interference that allows multiple FD radios/links to efficiently coexist. Consequently, we developed a novel MAC mechanism that enables concurrent operation of FD links while exploiting the unique advantages of FD radios (in learning radio medium at a much finer level than just carrier sensing and the ability to instantaneously adjust/adapt transmission behavior) to reduce network interference (that, in return, facilitates FD links' co-existence). The FD-MAC is fully distributed and converges to the unique NE whose efficacy is almost the same as that of the centralized algorithm. By leveraging SPs and SIS capability, the FD-MAC protocol significantly outperforms a CSMA-based design or the NOMA MAC with HD radios in terms of both throughput and energy/power efficiency.

The key idea behind exploiting SPs is inspired by the network duality [40] and the concept of spatial signature [19] [24]. Although network interference is often treated as colored noise then that gets whiten during the signal detection process, unlike random noise, network interference has its own structure. This structure is determined by the channel state information, signal precoding methods, and modulation. It can be "mined" for insights or intelligence to better align the transceiver's signal. This interference mining approach can be very helpful in facilitating concurrent transmissions in NOMA. Our future works will focus on investigating other applications of FD radios to NOMA. Additionally, it is also interesting to evaluate the energy efficiency of FD radios for short transmission distances, where the circuitry energy consumption is quite significant compared to transmission energy.

REFERENCES

[1] D. Nguyen and M. Krunz, "Be responsible: A Novel Communications Scheme for Full-duplex MIMO radios," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM)*, April 2015, pp. 1733–1741.

[2] W. Paper, "The power of wireless cloud," University of Melbourne," Tech. Rep., 2013. [Online]. Available: http://www.ceet.unimelb.edu.au/publications/downloads/ceet-white-paper-wireless-cloud.pdf

[3] J. I. Choi, M. Jain, K. Srinivasan, P. Levis, and S. Katti, "Achieving single channel, full duplex wireless communication," in *Proceedings of MobiCom Conference*, 2010, pp. 1–12.

[4] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proceedings of MobiCom Conference*, 2011, pp. 301–312.

[5] A. Sabharwal, P. Schniter, D. Guo, D. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 9, pp. 1637–1652, Sept 2014.

[6] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," in *Proceedings of SIGCOMM Conference*, 2013.

[7] D. Bharadia and S. Katti, "Full duplex MIMO radios," in *Proceedings of the 11th USENIX Symposium on Networked Systems Design and Implementation (NSDI)*, 2014.

[8] Y. Yang, B. Chen, K. Srinivasan, and N. B. Shroff, "Characterizing the achievable throughput in wireless networks with two active rf chains," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[9] X. Xie and X. Zhang, "Does full-duplex double the capacity of wireless networks," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[10] D. N. Nguyen, M. Krunz, and S. Hanly, "On the throughput of full-duplex MIMO in the multi-link case," in *Proceedings of the IEEE WiOpt Conference*, 2014.

[11] R. K. Mungara, I. Thibault, and A. Lozano, "Full-duplex MIMO in cellular networks: System-level performance," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3124–3137, 2017.

[12] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for Full Duplex Multiuser MIMO Systems: Spectral and Energy Efficiency Maximization," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4038–4050, 2013.

[13] M. Maso, C.-F. Liu, C.-H. Lee, T. Quek, and L. Cardoso, "Energy-Recycling Full-Duplex Radios for Next-Generation Networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, pp. 2948–2962, 2015.

[14] Z. Wei, X. Zhu, S. Sun, Y. Huang, L. Dong, and Y. Jiang, "Full-Duplex Versus Half-Duplex Amplify-and-Forward Relaying: Which is More Energy Efficient in 60-GHz Dual-Hop Indoor Wireless Systems?" *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 12, pp. 2936–2947, 2015.

[15] P. Aquilina, A. C. Cirik, and T. Ratnarajah, "Weighted sum rate maximization in full-duplex multi-user multi-cell MIMO networks," *IEEE Transactions on Communications*, vol. 65, no. 4, pp. 1590–1608, 2017.

[16] Y. Zhang, L. Lazos, K. Chen, B. Hu, and S. Shivaramaiah, "FD-MMAC: combating multi-channel hidden and exposed terminals using a single transceiver," in *Proceedings of the IEEE INFOCOM Conference*, 2014.

[17] W. Afifi and M. Krunz, "Exploiting self-interference suppression for improved spectrum awareness/efficiency in cognitive radio systems," in *Proceedings of the IEEE INFOCOM Conference*, 2013.

[18] M. Amjad, F. Akhtar, M. H. Rehmani, M. Reisslein, and T. Umer, "Full-duplex communication in cognitive radio networks: A survey," *IEEE Communications Surveys Tutorials*, no. 99, pp. 1–15, 2017.

[19] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, May 2005.

[20] F. Rashid-Farrokhi, L. Tassiulas, and K. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Transactions on Communications*, vol. 46, no. 10, pp. 1313–1324, Oct 1998.

[21] D. Nguyen, L. N. Tran, P. Pirinen, and M. Latva-aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Transactions on Signal Processing*, vol. 61, no. 16, pp. 4038–4050, 2013.

[22] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Energy efficient beamforming design for full-duplex MIMO interference channels," in *2017 IEEE International Conference on Communications (ICC)*, 2017, pp. 1–6.

[23] R. K. Mungara, I. Thibault, and A. Lozano, "Full-duplex MIMO in cellular networks: System-level performance," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 3124–3137, 2017.

[24] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 7, pp. 1341–1347, Sep 1995.

[25] N. Bambos, "Toward power-sensitive network architectures in wireless communications: concepts, issues, and design aspects," *IEEE Personal Communications*, vol. 5, no. 3, pp. 50–59, Jun. 1998.

[26] D. Nguyen and M. Krunz, "Power minimization in mimo cognitive networks using beamforming games," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 916–925, May 2013.

[27] A. Auslender and M. Teboulle, *Asymptotic Cones and Functions in Optimization and Variational Inequalities*. Springer, 2003.

[28] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Processing Magazine*, pp. 35–47, May 2010.

[29] D. Hoang and R. Iltis, "Noncooperative eigencoding for MIMO ad hoc networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 2, pp. 865–869, 2008.

[30] M. J. Osborne, *An Introduction to Game Theory*. Oxford University Press, 2004.

[31] A. Berman and R. Plemmons, "Nonnegative matrices in the mathematical sciences," *Society for Industrial Mathematics (SIAM)*, 1987.

[32] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1990.

[34] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1989.

[35] J. Wang, G. Scutari, and D. P. Palomar, "Robust MIMO cognitive radio via game theory," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1183–1201, March 2011.

[36] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1995.

[37] Y. Li, B. Bakkaloglu, and C. Chakrabarti, "A system level energy model and energy-quality evaluation for integrated transceiver front-ends," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 15, no. 1, pp. 90–103, Jan 2007.

[38] J. Zhou, T. H. Chuang, T. Dinc, and H. Krishnaswamy, "Integrated wideband self-interference cancellation in the RF domain for FDD and full-duplex wireless," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 12, pp. 3015–3031, 2015.

[39] D. Nguyen and M. Krunz, "Power-efficient spatial multiplexing for multi-antenna manets," in *Proc. of the ICC Conf.*, 2012.

[40] P. Viswanath and D. N. Tse, "Sum capacity of the vector gaussian broadcast channel and uplink-downlink duality," *IEEE Transaction on Information Theory*, vol. 49, no. 8, pp. 1912–1921, Sep. 2003.

[41] F. Facchinei and J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer-Verlag New York, Inc, 2003.

APPENDIX A
PROOF OF LEMMA 1

We first introduce the following proposition:

Proposition 1: Let $P^* = \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)$ and $c_i^{l*} = d_i^l$ be the transmit power by radio i_r and the corresponding throughput received by radio i_l after updating its precoder as $\mathbf{G}_i^r$. This precoder $\mathbf{G}_i^r$ must be a solution to:

$$\begin{aligned} & \underset{\{\mathbf{G}_i^r\}}{\text{maximize}} \;\; c_i^l \\ & \text{s.t.} \quad \text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r) \leq P^*. \end{aligned} \tag{28}$$

$$M+g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}\mathbf{H}_{ii}^{ll\,H}\mathbf{G}_i^l\mathbf{G}_i^{l\,H})+\sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}\mathbf{G}_j^{lH}\mathbf{G}_j^l)+\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^{rH}\mathbf{G}_j^r))\geq M|\mathbf{H}_{ii}^{lr}\mathbf{V}_i^l\mathbf{W}_i^{l-1}\mathbf{V}_i^{lH}\mathbf{H}_{ii}^{lrH}|^{\frac{1}{M}} \tag{30a}$$

$$M+g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}\mathbf{H}_{ii}^{ll\,H}\mathbf{G}_i^l\mathbf{G}_i^{l\,H})+\sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll}\mathbf{G}_j^{lH}\mathbf{G}_j^l)+\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr}\mathbf{G}_j^{rH}\mathbf{G}_j^r))\geq M|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}\frac{1}{\text{eig}_{\max}(\mathbf{W}_i^l)} \tag{30b}$$

$$M+g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}\mathbf{H}_{ii}^{ll\,H})\text{tr}(\mathbf{G}_i^l\mathbf{G}_i^{l\,H})+\sum_{j=1|j\neq i}^{N}(\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l)+\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r))\geq M|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}\frac{1}{\text{eig}_{\max}(\mathbf{W}_i^l)} \tag{30c}$$

$$\text{eig}_{\max}(\mathbf{W}_i^l)\geq\frac{|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1+g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l)+\sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l)+\frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r))} \tag{30d}$$

$$d_i^l\geq\log(1+\frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1+g_{sis}\frac{\text{tr}(\mathbf{H}_{ii}^{llH}\mathbf{H}_{ii}^{ll})}{M}\text{tr}(\mathbf{G}_i^{lH}\mathbf{G}_i^l)+\sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r)+\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l))}) \tag{30e}$$

Proof: If $\mathbf{G}_i^r$ is not a solution to problem (28) (given it can attains rate d_i^l as being assumed) then there exists a precoder $\bar{\mathbf{G}}_i^r$ that requires at most power of P^* but achieves a rate $\bar{c}_i^l > d_i^l$. In other words, it is possible for radio i_r to reduce its transmit power to achieve a rate of d_i^l. This contradicts to the fact that $\mathbf{G}_i^r$ is the optimal solution (or best response) of i_r (i.e., transmit at lower power). Thus, $\mathbf{G}_i^r$ must be a solution of (28). □

From the definition of $\mathbb{P}_{feasible}(\boldsymbol{d})$ (7), we have the following inequality $\forall i$ radios:

$$d_i^l\geq\log|\mathbf{I}+\mathbf{G}_i^{rH}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}\mathbf{G}_i^r| \tag{29a}$$
$$\geq\log\left(1+\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr})\right) \tag{29b}$$

where the (29a) comes from the Proposition 1 and the RHS of (29b) is a lower-bound of the rate c_i^l in problem (28), obtained by allocating all power $\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)$ on the subchannel $\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr})$ and zero power on others.

Besides, let $\mathbf{V}_{ii}^l$ be the unitary matrix that diagonalizes matrix $\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}$ and the diagonal matrix $\mathbf{W}_i^l$ contain eigenvalues of matrix $\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}$, we have:

$$\mathbf{V}_{ii}^{lH}\mathbf{H}_{ii}^{lrH}\mathbf{Q}_i^{l^{-1}}\mathbf{H}_{ii}^{lr}\mathbf{V}_{ii}^l=\mathbf{W}_i^l \tag{30a}$$
$$\text{tr}(\mathbf{Q}_i^l)=\text{tr}(\mathbf{H}_{ii}^{lr}\mathbf{V}_{ii}^l\mathbf{W}_i^{l-1}\mathbf{V}_{ii}^{lH}\mathbf{H}_{ii}^{lrH}) \tag{30b}$$

Then we have equations (30a),(30b),(30c),(30d). (30a) follows by recalling the noise-plus-interference covariance matrix on the LHS of (30b) and applying the identity $\frac{\text{tr}(\mathbf{A})}{n}\geq|\mathbf{A}|^{1/n}$ [33] (for any $n\times n$ positive semi-definite matrix $\mathbf{A}$) to the RHS of (30b). (30c) follows from applying the identity $\text{tr}(\mathbf{AB})\leq\text{tr}(\mathbf{A})\text{tr}(\mathbf{B})$ to the LHS of (30b).

From inequalities (30d) and (29b), we get (30e). We then have:

$$(2^{d_i^l}-1)\geq\mathbf{\Gamma}(2i-1,:)\times \tag{31a}$$
$$[\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l),\text{tr}(\mathbf{G}_1^{rH}\mathbf{G}_1^r),\ldots,\text{tr}(\mathbf{G}_N^{lH}\mathbf{G}_N^l),\text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T \tag{31b}$$
$$\mathbf{\Gamma}^{-1}\times[2^{d_1^l}-1,2^{d_1^r}-1,\ldots,2^{d_N^l}-1,2^{d_N^r}-1]^T\geq \tag{31c}$$
$$[\text{tr}(\mathbf{G}_1^{lH}\mathbf{G}_1^l),\text{tr}(\mathbf{G}_1^{rH}\mathbf{G}_1^r),\ldots,\text{tr}(\mathbf{G}_N^{lH}\mathbf{G}_N^l),\text{tr}(\mathbf{G}_N^{rH}\mathbf{G}_N^r)]^T$$

where the inequality (31c) comes from the assumption $\mathbf{\Gamma}$ is a P-matrix, hence invertible [31].

Hence the RHS in (31c) is bounded or rate demands can be fullfilled with a bounded power allocation vector $\boldsymbol{p}$. In other words, $\mathbb{P}_{feasible}(\boldsymbol{d})$ contains at least one bounded $\boldsymbol{p}$. □

APPENDIX B
PROOF OF LEMMA 2

For $\boldsymbol{f}\in\mathbb{Q}^{asymp}(\boldsymbol{d})$, by the definition of limit directions, there exists sequences $\{\boldsymbol{p_n}\}$ and $\{\nu_n\}$. Consequently, we have equations (32) where (32b), (32c), and (32d) follow from (30d), (11), and the definition of $\boldsymbol{d}$, respectively. □

APPENDIX C
PROOF OF THEOREM 3

Variational Inequality (VI) problem: [41] Given a subset $\mathbb{K}$ of the Euclidean n-dimensional space $\mathbb{R}^n$ and a mapping $F:\mathbb{K}\rightarrow\mathbb{R}^n$, a $\text{VI}(\mathbb{K},\mathbb{R}^n)$ problem is to find a vector $\boldsymbol{x^{opt}}\in\mathbb{K}$ so that:

$$(\boldsymbol{x}-\boldsymbol{x^{opt}})^TF(\boldsymbol{x^{opt}})\geq 0,\quad\forall\boldsymbol{x}\in\mathbb{K}. \tag{33}$$

If the set $\mathbb{K}$ has a Cartesian structure, i.e., $\mathbb{K}=\mathbb{K}_1\times\mathbb{K}_2\times\ldots\times\mathbb{K}_N$ (where $\mathbb{K}_i\in\mathbb{R}^{n_i}$ and $\sum_{i=1}^{N}n_i=n$), we have the following theorem regarding the existence and uniqueness of a solution to the above VI problem (Proposition 3.5.10 in [41]).

Theorem 5: If set $\mathbb{K}$ has a Cartesian structure, the $\text{VI}(\mathbb{K},\mathbb{R}^n)$ problem has a unique solution $\boldsymbol{x}^{opt}$ provided $\mathbb{K}_i$ is closed and convex and F is a continuous uniformly-P function, i.e, there exists a positive constant α such that:

$$\max_{\{1\leq i\leq N\}}(\boldsymbol{x}_i-\boldsymbol{x}'_i)^T(F_i(\boldsymbol{x}_i)-F_i(\boldsymbol{x}'_i))\geq\alpha\|\boldsymbol{x}-\boldsymbol{x}'\|^2,\quad\forall\boldsymbol{x},\boldsymbol{x}'\in\mathbb{K}. \tag{34}$$

To cast the game (6) as a VI problem, we use the $vec()$ operator in (25) to map the complex matrix in (6) to the Euclidean domain, by stacking columns (from left to right) of an $m\times n$ matrix to form an $mn\times 1$ vector. The gradient of a matrix function (.) w.r.t $\mathbf{G}_i^r$ is in (26).

$$
\begin{aligned}
d_i^l = c_i^l(\boldsymbol{p}_n) &\ge \log\left(1+\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)\text{eig}_{\max}(\mathbf{H}_{ii}^{lrH}{\mathbf{Q}_i^l}^{-1}\mathbf{H}_{ii}^{lr})\right) && \text{(32a)}\\
&\ge \log(1+\frac{\text{tr}(\mathbf{G}_i^{rH}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{1+g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H)\text{tr}(\mathbf{G}_i^l{\mathbf{G}_i^l}^H)+\sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\mathbf{G}_j^l)+\frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\mathbf{G}_j^r))}) && \text{(32b)}\\
&= \log(1+\frac{\text{tr}(\mathbf{G}_i^{rH}\frac{1}{\nu_n}\mathbf{G}_i^r)|\mathbf{H}_{ii}^{lrH}\mathbf{H}_{ii}^{lr}|^{\frac{1}{M}}}{\frac{1}{\nu_n}+g_{sis}\text{tr}(\mathbf{H}_{ii}^{ll}{\mathbf{H}_{ii}^{ll}}^H)\text{tr}(\mathbf{G}_i^l\frac{1}{\nu_n}{\mathbf{G}_i^l}^H)+\sum_{j=1|j\neq i}^{N}(\frac{\text{tr}(\mathbf{H}_{ij}^{llH}\mathbf{H}_{ij}^{ll})}{M}\text{tr}(\mathbf{G}_j^{lH}\frac{1}{\nu_n}\mathbf{G}_j^l)+\frac{\text{tr}(\mathbf{H}_{ij}^{lrH}\mathbf{H}_{ij}^{lr})}{M}\text{tr}(\mathbf{G}_j^{rH}\frac{1}{\nu_n}\mathbf{G}_j^r))}) && \text{(32c)}\\
&\overset{n\to+\infty}{=} {c'}_i^l(\boldsymbol{f}) \qquad \forall i_l, i_r. && \text{(32d)}
\end{aligned}
$$

If the condition in Theorem 2 holds, the strategic space of player i_r, denoted by $\mathcal{G}_i^r \in \mathbb{C}^{M\times M}$, is nonempty. Moreover, it can be verified that $\mathcal{G}_i^r$ is convex and bounded. Hence, problem (6) is convex. The following inequality captures the necessary (and also the sufficient) condition for strategy $\hat{\mathbf{G}}_i^r$ to be the optimal response:

$$(\mathbf{G}_i^r - \hat{\mathbf{G}}_i^r) \bullet \nabla U_i^r \ge 0 \quad \forall \mathbf{G}_i^r \in \mathcal{G}_i \tag{35}$$

where $\mathbf{A} \bullet \mathbf{B} \overset{\text{def}}{=} \text{vec}(\mathbf{A})^T\text{vec}(\mathbf{B})$ and $U_i^r \overset{\text{def}}{=} \text{tr}(\mathbf{G}_i^r{\mathbf{Q}_i^r}^T\mathbf{G}_i^{rH})$

Define $\mathcal{G} \overset{\text{def}}{=} \mathcal{G}_1^l \times \mathcal{G}_1^r \dots \mathcal{G}_N^l \times \mathcal{G}_N^r$ and $F \overset{\text{def}}{=} F_1^l \times F_1^r \dots \times F_N^l \times F_N^r$ with $F_i^r \overset{\text{def}}{=} \nabla U_i^r$. By comparing (35) with the definition of a VI problem, the set $\hat{\mathbf{G}} \overset{\text{def}}{=} [\hat{\mathbf{G}}_1^l \times \hat{\mathbf{G}}_1^r \dots \times \hat{\mathbf{G}}_N^l \times \hat{\mathbf{G}}_N^r]$ is a NE of the game (6) iff $\hat{\mathbf{G}}$ is a solution of the VI$(\mathcal{G}, F)$ problem. Note that the existence of a NE guarantees that $\mathbb{K}_u$ is closed and convex. The next step is to show that given a NE existence, F is a continuous uniformly-P function.

Let $\mathbf{G} \overset{\text{def}}{=} [\mathbf{G}_1^l \times \mathbf{G}_1^r \dots \times \mathbf{G}_N^l \times \mathbf{G}_N^r]$ and $\breve{\mathbf{G}} \overset{\text{def}}{=} [\breve{\mathbf{G}}_1^l \times \breve{\mathbf{G}}_1^r \dots \times \breve{\mathbf{G}}_N^l \times \breve{\mathbf{G}}_N^r]$ be two different strategy sets of the strategic space $\mathcal{G}$ of the game (6), then:

$$F_i^r(\breve{\mathbf{G}}_i^r) = [\mathbf{S}_i^r + \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{rrH}\breve{\mathbf{G}}_i^r{\breve{\mathbf{G}}_i^r}^H\mathbf{H}_{ii}^{rr}]\breve{\mathbf{G}}_i^r \tag{36}$$

Consequently:

$$\text{vec}(\breve{\mathbf{G}}_i^r - \mathbf{G}_i^r)^T\text{vec}(F_i(\breve{\mathbf{G}}_i^r) - F_i(\mathbf{G}_i^r)) \ge \alpha_i^r||\text{vec}((\breve{\mathbf{G}}_i^r - \mathbf{G}_i^r))||^2 \tag{37}$$

where the above inequality follows from the fact that $||\mathbf{A}\boldsymbol{a}|| \ge \text{eig}_{\min}(\mathbf{A})||\boldsymbol{a}||$ and $\alpha_i^r \overset{\text{def}}{=} \text{eig}_{\min}[\mathbf{S}_i^r + \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{rrH}\breve{\mathbf{G}}_i^r{\breve{\mathbf{G}}_i^r}^H\mathbf{H}_{ii}^{rr}]$

We then have:

$$\sum_{i=1}^{N}\text{vec}(\breve{\mathbf{G}}_i^r - \mathbf{G}_i^r)^T\text{vec}(F_i(\breve{\mathbf{G}}_i^r) - F_i(\mathbf{G}_i^r)) \ge \alpha_i^r\sum_{i=1}^{N}||\text{vec}((\breve{\mathbf{G}}_i^r - \mathbf{G}_i^r))||^2 \tag{38}$$

Similarly:

$$\sum_{i=1}^{N}\text{vec}(\breve{\mathbf{G}}_i^l - \mathbf{G}_i^l)^T\text{vec}(F_i(\breve{\mathbf{G}}_i^l) - F_i(\mathbf{G}_i^l)) \ge \alpha_i^l\sum_{i=1}^{N}||\text{vec}((\breve{\mathbf{G}}_i^l - \mathbf{G}_i^l))||^2 \tag{39}$$

where $\alpha_i^l \overset{\text{def}}{=} \text{eig}_{\min}[\mathbf{S}_i^l + \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{llH}\breve{\mathbf{G}}_i^l{\breve{\mathbf{G}}_i^l}^H\mathbf{H}_{ii}^{ll}]$

Summing up the above two inequalities for all $i = 1, ..., N$ and recalling the triangle inequality, we have:

$$
\begin{aligned}
\max_{\{1\le i\le N\}}\{&\text{vec}(\breve{\mathbf{G}}_i^l - \mathbf{G}_i^l)^T\text{vec}(F_i(\breve{\mathbf{G}}_i^l) - F_i(\mathbf{G}_i^l)),\\
&\text{vec}(\breve{\mathbf{G}}_i^r - \mathbf{G}_i^r)^T\text{vec}(F_i(\breve{\mathbf{G}}_i^r) - F_i(\mathbf{G}_i^r))\}\\
&\ge \alpha||\text{vec}((\breve{\mathbf{G}} - \mathbf{G}))||^2
\end{aligned} \tag{40}
$$

where $\alpha = \min_{\{\le i\le N\}}\{\frac{\alpha_i^r}{2N}, \frac{\alpha_i^l}{2N}\}$.

Since $[\mathbf{S}_i^r + \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{rrH}\breve{\mathbf{G}}_i^r{\breve{\mathbf{G}}_i^r}^H\mathbf{H}_{ii}^{rr}]$ and $[\mathbf{S}_i^l + \mathbf{I} + g_{sis}\mathbf{H}_{ii}^{llH}\breve{\mathbf{G}}_i^l{\breve{\mathbf{G}}_i^l}^H\mathbf{H}_{ii}^{ll}]$ are positive definite, so are $\alpha_i^r, \alpha_i^l, \alpha$. Hence, the above mapping F is a continuous uniformly-P function. The VI$(\mathcal{G}, F)$ problem or game (6) has a unique NE. □

Diep N. Nguyen is a faculty member of the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS). He received M.E. and Ph.D. in Electrical and Computer Engineering from the University of California San Diego (UCSD) and The University of Arizona (UA), respectively. Before joining UTS, he was a DECRA Research Fellow at Macquarie University, a member of technical staff at Broadcom (California), ARCON Corporation (Boston), consulting the Federal Administration of Aviation on turning detection of UAVs and aircraft, US Air Force Research Lab on anti-jamming. He has received several awards from LG Electronics, University of California, San Diego, The University of Arizona, US National Science Foundation, Australian Research Council. His recent research interests are in the areas of computer networking, wireless communications, and machine learning application, with emphasis on systems' performance and security/privacy.

Marwan Krunz is the Kenneth VonBehren Endowed Professor in the ECE Department at the University of Arizona. He is also an affiliated faculty member of the University of Technology Sydney (UTS). He co-directs the Broadband Wireless Access and Applications Center, a multi-university industry-focused NSF center that includes affiliates from industry and government labs. He previously served as the UA site director for Connection One, an NSF IUCRC that focuses on wireless communication circuits and systems. In 2010, Dr. Krunz was a Visiting Chair of Excellence at the University of Carlos III de Madrid. He previously held visiting research positions at UTS, INRIA-Sophia Antipolis, HP Labs, University of Paris VI, University of Paris V, University of Jordan, and US West Advanced Technologies. Dr. Krunzs research interests lie in the areas of wireless communications and networking, with emphasis on resource management, adaptive protocols, and security issues. He has published more than 250 journal articles and peer-reviewed conference papers, and is a co-inventor on several US patents. He is an IEEE Fellow, an Arizona Engineering Faculty Fellow (2011-2014), and an IEEE Communications Society Distinguished Lecturer (2013 and 2014). He was the recipient of the 2012 IEEE TCCC Outstanding Service Award. He received the NSF CAREER award in 1998. He currently serves as the Editor-in-Chief for the IEEE Transactions on Mobile Computing (TMC). He previously served on the editorial boards for the IEEE Transactions on Cognitive Communications and Networks, IEEE/ACM Transactions on Networking, IEEE TMC, IEEE Transactions on Network and Service Management, Computer Communications Journal, and IEEE Communications Interactive Magazine. He was the general vice-chair for WiOpt 2016 and general co-chair for WiSec12. He was the TPC chair for WCNC 2016 (Networking Track), INFOCOM04, SECON05, WoWMoM06, and Hot Interconnects 9. He has served and continues to serve on the steering and advisory committees of numerous conferences and on the panels of several funding agencies. He was a keynote speaker, an invited panelist, and a tutorial presenter at numerous international conferences. See http://www2.engr.arizona.edu/ krunz/ for more details.

Eryk Dutkiewicz received his B.E. degree in Electrical and Electronic Engineering from the University of Adelaide in 1988, his M.Sc. degree in Applied Mathematics from the University of Adelaide in 1992 and his PhD in Telecommunications from the University of Wollongong in 1996. His industry experience includes management of the Wireless Research Laboratory at Motorola in early 2000s. He has co-authored 17 US patents filings and published over 280 journal and conference papers. He is currently the Head of School of Electrical and Data Engineering at the University of Technology Sydney, Australia. He has held visiting professorial appointments at several institutions including the Chinese Academy of Sciences, Shanghai JiaoTong University and Macquarie University. His current research interests cover 5G networks and medical body area networks.