Secure Transmissions Using Artificial Noise in MIMO Wiretap Interference Channel: A Game Theoretic Approach

Peyman Siyari[1], Marwan Krunz[1], and Diep N. Nguyen[2]
[1]Department of Electrical and Computer Engineering, University of Arizona, USA
[2]Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
{psiyari, krunz}@email.arizona.edu, diep.nguyen@uts.edu.au

Technical Report
TR-UA-ECE-2016-2
Last Update: December 21, 2016

Abstract

We consider joint optimization of artificial noise (AN) and information signals in a MIMO wiretap interference network, wherein the transmission of each link may be overheard by several MIMO-capable eavesdroppers. Each information signal is accompanied with AN, generated by the same user to confuse nearby eavesdroppers. Using a noncooperative game, a distributed optimization mechanism is proposed to maximize the secrecy rate of each link. The decision variables here are the covariance matrices for the information signals and ANs. However, the nonconvexity of each link's optimization problem (i.e., best response) makes conventional convex games inapplicable, even to find whether a Nash Equilibrium (NE) exists. To tackle this issue, we analyze the proposed game using a relaxed equilibrium concept, called *quasi-Nash equilibrium* (QNE). Under a constraint qualification condition for each player's problem, the set of QNEs includes the NE of the proposed game. We also derive the conditions for the existence and uniqueness of the resulting QNE. It turns out that the uniqueness conditions are too restrictive, and do not always hold in typical network scenarios. Thus, the proposed game often has multiple QNEs, and convergence to a QNE is not always guaranteed. To overcome these issues, we modify the utility functions of the players by adding several specific terms to each utility function. The modified game converges to a QNE even when multiple QNEs exist. Furthermore, players have the ability to select a desired QNE that optimizes a given social objective (e.g., sum-rate or secrecy sum-rate). Depending on the chosen objective, the amount of signaling overhead as well as the performance of resulting QNE can be controlled. Simulations show that not only we can guarantee the convergence to a QNE, but also due to the QNE selection mechanism, we can achieve a significant improvement in terms of secrecy sum-rate and power efficiency, especially in dense networks.

Index Terms

Wiretap interference network, friendly jamming, quasi-Nash equilibrium, NE selection, nonconvex games.

I. Introduction

PHYSICAL-layer (PHY-layer) security provides a cost-efficient alternative to cryptographic methods in scenarios where the use of the latter is either impractical or expensive. One of the common settings for PHY-layer security is the wiretap channel. In this channel, a node (Alice) wishes to transmit messages securely to a legitimate receiver (Bob) in the presence of one or more eavesdroppers (Eve). Most PHY-layer security techniques for the wiretap channel are based on an information-theoretic definition of security, namely, the *secrecy capacity*, defined as the largest amount of information that can be confidentially communicated between Alice and Bob [1].

Over the last decade, several PHY-layer security techniques have been proposed. Some of these techniques rely on the use of artificial noise (AN) as a friendly jamming (FJ) signal [2]. In this method, Alice uses multiple antennas to generate an FJ signal along with the information signal, increasing the

interference at Eve but without affecting Bob. The authors in [2] proposed a simple version of this technique, which relies on MIMO zero-forcing to ensure that the FJ signal falls in the null-space of the channel between Alice and Bob. The interest in using AN for a single link is driven by pragmatic considerations, and not necessarily due to its optimality. In fact, it was shown in [3] that in one-Eve scenario, the optimal approach for securing a single link with the knowledge of Eve's location is not to use AN. Complementing the classic AN approach in [2], which relies on transmitting the AN in the null-space of the legitimate channel, it was shown in [4] that adding AN to both the legitimate channel and its null-space can further improve the secrecy rate of a link. In the case of multiple eavesdroppers, it was shown in [5] that the use of AN can significantly improve the secrecy rate compared to the case when AN is not used.

In a multi-link scenario, where several transmitters wish to convey their messages simultaneously to several legitimate receivers (see Fig. 1), the FJ signal of each transmitter must be designed to not interfere with other unintended (but legitimate) receivers in the network. This can be quite challenging when only limited or no coordination is possible between links. Therefore, providing PHY-layer secrecy has to be done in a distributed yet noninterfering manner. Interference management for PHY-layer security involves two conflicting factors. On the one hand, the AN from one transmitter degrades the respective information signals at unintended (but legitimate) receivers. On the other hand, AN also increases the interference at eavesdroppers, and is hence useful in terms of improving the security of the communications. The idea of using interference in networks to provide secrecy was first discussed in [6]. Several subsequent works exploited this idea in other applications, such as the coexistence of different protocols on the same channel. For instance, the authors in [7] considered a two-link SISO interference network, which resembles a coexistence scenario. They showed that with a careful power control design for both links, one link can assist the other in providing a rate demand guarantee as well as secure the transmission by increasing interference on a single-antenna eavesdropper. For the case of two transmitter-receiver-eavesdropper triples, the authors in [8] proposed a cooperative beamforming approach to achieve maximum secure degree of freedom for both users. In fact, given the knowledge of co-channel interference at the receivers, a cooperative transmission alignment scheme between transmitters is established such that their respective receivers will get interference-free signals and the eavesdropper corresponding to each link will receive interference.

In this paper, we consider a peer-to-peer multi-link interference network in which the transmission on each link can be overheard by several external eavesdroppers. Perhaps the closest works to our scenario are [7] and [8]. Apart from the fact that both of these works consider only a two-user scenario, which limits their applicability, in [7] one of the users generates only interference to provide PHY-layer security for the other user, so providing the PHY-layer secrecy of the former user is overlooked. In contrast, our work provides PHY-layer security for all users. Moreover, although the work in [8] considers providing secrecy for both users (in a slightly different network than the one we consider), it requires a significant amount of signaling (i.e., coordination) between the two users. In this paper, we limit the amount of coordination as much as possible.

In our system model, we assume that the transmission of each information signal is accompanied with AN. Each node in the network is equipped with multiple antennas. Our goal is to design a framework through which the co-channel interference at each legitimate receiver is minimized while the aggregate interference at external eavesdroppers remains high. Because nodes cannot cooperate with each other in our settings, each link independently tries to maximize its secrecy rate by designing the covariance matrices (essentially, the precoders) of its information signal and AN. This independent secrecy optimization can be modeled via noncooperative game theory. Specifically, we design a game-theoretic framework in which the utility of each player (i.e., link) is his secrecy rate, and the player's strategy is to optimize the covariance matrices of information signal and AN. It turns out that finding the best response of each link requires solving a nonconvex optimization problem. Thus, the existence of a Nash Equilibrium (NE) cannot be proved using traditional concepts of convex (concave) games [9]. Instead, we study the proposed game based on a relaxed equilibrium concept called *quasi-Nash equilibrium* (QNE) [10]. A QNE is a solution of

a variational inequality (VI) [11] obtained under the K.K.T optimality conditions of the players' problems. We show that under a constraint qualification (CQ) condition for each player's problem, the set of QNEs also includes the NE. Sufficient conditions for the existence and uniqueness of the resulting QNE are provided. Then, an iterative algorithm is proposed to achieve the unique QNE.

Despite their attractiveness in terms of not requiring link coordination, the (Q)NEs of a purely noncooperative game are often inefficient in terms of the achievable sum-utility (i.e., secrecy sum-rate). Furthermore, the conditions that guarantee the uniqueness of the QNE are dependent on the channel gains between links. The random nature of channel gains greatly reduces the possibility of having a unique QNE, which further limits the effectiveness of the proposed noncooperative game. More specifically, the convergence to a QNE cannot be always achieved. This forces the links to terminate their iterative optimizations at some point, resulting in a low secrecy sum-rate. To overcome this issue, we introduce several modifications to the proposed game. Every modification appears as the addition of a term in the utility function of each player. These modifications allow us to not only guarantee the convergence of the game, but also give links the ability to selectively converge to a specific QNE among multiple QNEs. Selecting a particular QNE is done based on how much it satisfies a particular design criterion. We propose three possibilities for QNE selection, each providing different benefits and requiring a different amounts of communication overhead. The proposed QNE selection algorithm can improve the performance of the formerly proposed noncooperative game while keeping the communication overhead reasonably low.

The concept of QNE has been recently used in [12] in sum-rate maximization in cognitive radio users. However, no effort has been made to improve the performance of achieved QNEs. The work in [13] also considers the use of QNEs to jointly optimize the sensing and power allocation of cognitive radio users in the presence of primary users. Although in this work some improvements have been made on the performance of the resulting QNEs, they are specific to cognitive radios and thus not extendable to other networks. The framework we propose can be generalized to any similarly structured game. Overall, the contributions of this paper are as follows:

- We propose a noncooperative game to model PHY-layer secrecy optimization in a multi-link MIMO wiretap interference network. Due to the nonconvexity of each player's optimization problem, the analysis of equilibria is done through the concept of QNE. We show that the set of QNEs includes NE as well.
- Because many network scenarios may involve multiple QNEs, the purely noncooperative games do not always guarantee the convergence to a unique QNE. Hence, we introduce several modifications to the proposed game to guarantee the convergence to a QNE. The modifications appear as the additional terms in the utility function of the players and keep the distributed nature of the noncooperative game.
- We show that the modified game allows users to select a QNE among multiple QNEs according to a design criterion. QNE selection makes it possible to improve the resulting secrecy sum-rate of the modified game compared to a purely noncooperative game.
- We show that the freedom in choosing the design criterion gives a degree of flexibility to the modified game. We propose three different choices for the design criterion, each of which requiring a different level of coordination between links and offering a different amount secrecy sum-rate improvement.

The rest of this paper is organized as follows. In Section II, we introduce the system model. In Section III, we formulate the optimization of information signal and AN as a noncooperative game. The conditions for the existence and uniqueness of the QNE are established in Section IV. In Section V, we modify the proposed noncooperative game, and introduce the theoretical aspects of our QNE selection method. In Section VI an algorithm that implements the QNE selection is given and practical considerations are discussed. We present a centralized algorithm as a measure of efficiency of our proposed game in Section VII. In Section VIII simulation results assess the performance of our algorithms. Finally, Section IX concludes the paper.

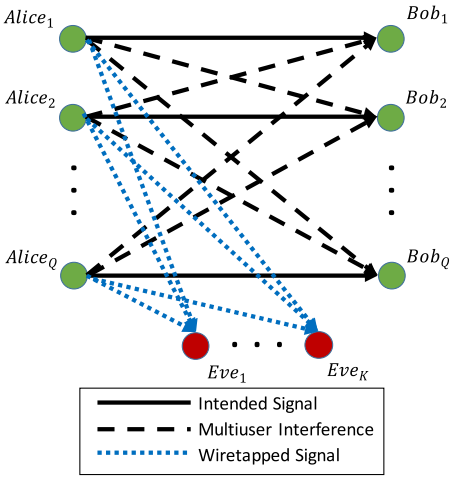


Fig. 1: System Model

II. System Model

Consider the network shown in Fig. 1, where Q transmitters, $Q > 1$, communicate with Q corresponding receivers. The qth transmitter is equipped with N_{T_q} antennas, $q = 1, \ldots, Q$. The qth receiver has N_{R_q} antennas, $q = 1, \ldots, Q$. The link between each transmit-receive (Alice-Bob) pair may experience interference from the other $Q-1$ links. There are K noncolluding Eves overhearing the communications. The kth Eve, $k = 1, \ldots, K$, has $N_{e,k}$ receive antennas[1]. The received signal at the qth receiver, $\mathbf{y}_q$, is

$$\mathbf{y}_q = \tilde{\mathbf{H}}_{qq}\mathbf{u}_q + \sum_{\substack{r=1 \\ r\neq q}}^{Q} \tilde{\mathbf{H}}_{rq}\mathbf{u}_r + \mathbf{n}_q, \;\; q \in \mathbb{Q} \tag{1}$$

where $\tilde{\mathbf{H}}_{rq}$ ($\tilde{\mathbf{H}}_{qq}$) denotes the $N_{R_q} \times N_{T_r}$ ($N_{R_q} \times N_{T_q}$) channel matrix between the rth (qth) transmitter and qth receiver, $\mathbf{u}_q$ is the $N_{T_q} \times 1$ vector of transmitted signal from the qth transmitter, $\mathbf{n}_q$ is the $N_{R_q} \times 1$ vector of additive noise whose elements are i.i.d zero-mean circularly symmetric complex Gaussian distributed with unit variance, and $\mathbb{Q} \triangleq \{1, \ldots, Q\}$. The term $\sum_{\substack{r=1 \\ r\neq q}}^{Q} \tilde{\mathbf{H}}_{rq}\mathbf{u}_r$ is the multi-user interference (MUI). The received signal at the kth eavesdropper, $\mathbf{z}_k$, is expressed as

$$\mathbf{z}_k = \sum_{q=1}^{Q} \mathbf{G}_{qk}\mathbf{u}_q + \mathbf{n}_{e,k}, \;\; k \in \mathbb{K} \tag{2}$$

where $\mathbf{G}_{qk}$ is the $N_{e,k} \times N_{T_q}$ channel matrix between the qth transmitter and the kth eavesdropper, $\mathbf{n}_{e,k}$ is the $N_{e,k} \times 1$ vector of additive noise at the kth eavesdropper, and $\mathbb{K} \triangleq \{1, \ldots, K\}$. The transmitted signal $\mathbf{u}_q$ has the following form:

$$\mathbf{u}_q \triangleq \mathbf{s}_q + \mathbf{w}_q \tag{3}$$

where $\mathbf{s}_q$ is the information signal and $\mathbf{w}_q$ is the AN. We use the Gaussian codebook for the information signal and the Gaussian noise for the AN[2]. The matrices $\mathbf{\Sigma}_q$ and $\mathbf{W}_q$ indicate the covariance matrices of $\mathbf{s}_q$ and $\mathbf{w}_q$, respectively.

The qth link, $q \in \mathbb{Q}$, together with K eavesdroppers form a compound wiretap channel for which the

[1]The treatment can be easily extended to colluding eavesdroppers by combining the K Eves into one with $\sum_{k=1}^{K} N_{e,k}$ antennas.

[2]Other practical codebooks for the information signal (e.g., QAM) can be approximated to a Gaussian codebook with a capacity gap (see [14]).

achievable secrecy rate of the qth link is written as [15]:

$$R_q^{sec}(\mathbf{\Sigma}_q, \mathbf{W}_q) \triangleq C_q(\mathbf{\Sigma}_q, \mathbf{W}_q) - \max_{k\in\mathbb{K}} C_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q),\ q \in \mathbb{Q} \tag{4}$$

where $C_q(\mathbf{\Sigma}_q, \mathbf{W}_q)$ is the information rate and $C_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q)$ is the received rate at the kth eavesdropper, $k \in \mathbb{K}$, while eavesdropping on the qth link, $q \in \mathbb{Q}$. Specifically,

$$\begin{aligned} C_q(\mathbf{\Sigma}_q, \mathbf{W}_q) &\triangleq \ln\left|\mathbf{I} + \mathbf{M}_q^{-1}\mathbf{H}_{qq}\mathbf{\Sigma_q}\mathbf{H}_{qq}^H\right| = \\ &\ln\left|\mathbf{M}_q + \mathbf{H}_{qq}\mathbf{\Sigma_q}\mathbf{H}_{qq}^H\right| + \ln\left|\mathbf{M}_q^{-1}\right| \end{aligned} \tag{5}$$

where $\mathbf{M}_q \triangleq \mathbf{I} + \mathbf{H}_{qq}\mathbf{W}_q\mathbf{H}_{qq}^H + \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{H}_{rq}\left(\mathbf{\Sigma}_r + \mathbf{W}_r\right)\mathbf{H}_{rq}^H$ and

$$\begin{aligned} C_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q) &\triangleq \ln\left|\mathbf{I} + \mathbf{M}_{e,q,k}^{-1}\mathbf{G}_{qk}\mathbf{\Sigma_q}\mathbf{G}_{qk}^H\right| = \\ &\ln\left|\mathbf{M}_{e,q,k} + \mathbf{G}_{qk}\mathbf{\Sigma_q}\mathbf{G}_{qk}^H\right| + \ln\left|\mathbf{M}_{e,q,k}^{-1}\right| \end{aligned} \tag{6}$$

where $\mathbf{M}_{e,q,k} \triangleq \mathbf{I} + \mathbf{G}_{qk}\mathbf{W}_q\mathbf{G}_{qk}^H + \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{G}_{rk}\left(\mathbf{\Sigma}_r + \mathbf{W}_r\right)\mathbf{G}_{rk}^H$. The term $\mathbf{M}_q$ is the covariance matrix of received interference at the qth receiver and $\mathbf{M}_{e,q,k}$ is the covariance matrix of interference received at the kth eavesdropper while eavesdropping on the qth link[3]. Notice that both $\mathbf{M}_q$ and $\mathbf{M}_{e,q,k}$ include the information signal and AN of other $Q-1$ links. Furthermore, we require $\text{tr}(\mathbf{\Sigma}_q + \mathbf{W}_q) \leq P_q$ for all $q \in \mathbb{Q}$, where tr(.) is the trace operator and P_q is a positive value that represents the amount of power available (for both information and AN signals) at the qth transmitter.

III. Problem Formulation

We assume that the qth link, $q \in \mathbb{Q}$, optimizes its information and AN signals (through their covariance matrices $\mathbf{\Sigma}_q$ and $\mathbf{W}_q$) to maximize its own secrecy rate. The dynamics of such interaction between Q links can be modeled as a noncooperative game where each player (i.e., link) uses his best strategy to maximize his own utility (i.e., secrecy rate) given the strategies of other players. The best response of each player can be found by solving the following optimization problem

$$\begin{aligned} &\underset{\mathbf{\Sigma}_q, \mathbf{W}_q}{\text{maximize}}\ R_q^{sec}(\mathbf{\Sigma}_q, \mathbf{W}_q) \\ &\text{s.t.} \qquad (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q,\ q \in \mathbb{Q} \end{aligned} \tag{7}$$

where $\mathcal{F}_q \triangleq \{(\mathbf{\Sigma}_q, \mathbf{W}_q) | \text{tr}(\mathbf{\Sigma}_q + \mathbf{W}_q) \leq P_q,\ \mathbf{\Sigma}_q \succeq 0,\ \mathbf{W}_q \succeq 0\}$ is the set of all Hermitian matrices $(\mathbf{\Sigma}_q, \mathbf{W}_q)$ that are positive semi-definite (i.e., $\mathbf{\Sigma}_q \succeq 0,\ \mathbf{W}_q \succeq 0$) and meet the link's power constraint.

Unfortunately, problem (7) is a nonconvex optimization problem. In the remainder of this section, we aim to find a tractable solution for this problem. To that end, we first mention the following identity for a positive definite matrix $\mathbf{M}_q$ of size N_{R_q} [16, Example 3.23]:

$$\ln|\mathbf{M}_q^{-1}| = f(\mathbf{S}^*) = \max_{\mathbf{S}\in\mathbb{C}^{N_{R_q}\times N_{R_q}}, \mathbf{S}\succeq 0} f(\mathbf{S}) \tag{8}$$

where $f(\mathbf{S}) \triangleq -\text{tr}(\mathbf{S}\mathbf{M}_q) + \ln|\mathbf{S}| + N_{R_q}$ and $\mathbf{S}^* \triangleq \mathbf{M}_q^{-1}$ is the solution to the most RHS of (8). Applying the reformulation in (8) to the term $\ln|\mathbf{M}_q^{-1}|$ in (5) and $\ln\left|\mathbf{M}_{e,q,k} + \mathbf{G}_{qk}\mathbf{\Sigma_q}\mathbf{G}_{qk}^H\right|$ in (6), (7) can be rewritten as

$$\begin{aligned} &\underset{\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_q}{\text{maximize}}\ f_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^K), \\ &\text{s.t.} \qquad (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q,\ \mathbf{S}_{q,k} \succeq 0,\ q \in \mathbb{Q},\ k \in \{0\} \cup K \end{aligned} \tag{9}$$

[3]Specifically, while eavesdropping on a user, an eavesdropper is treating interference as additive (colored) noise.

where $\{\mathbf{S}_{q,k}\}_{k=0}^{K} = [\mathbf{S}_{q,0}^T, \ldots, \mathbf{S}_{q,K}^T]^T$, and

$$f_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K}) \triangleq \varphi_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,0}) - \max_{k\in\mathbb{K}} \varphi_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,k}) \tag{10a}$$

$$\varphi_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,0}) \triangleq -\text{tr}(\mathbf{S}_{q,0}\mathbf{M}_q) + \ln|\mathbf{S}_{q,0}| + N_{R_q} + \ln\left|\mathbf{M}_q + \mathbf{H}_{qq}\mathbf{\Sigma_q}\mathbf{H}_{qq}^H\right| \tag{10b}$$

$$\varphi_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,k}) \triangleq \text{tr}(\mathbf{S}_{q,k}(\mathbf{M}_{e,q,k} + \mathbf{G}_{qk}\mathbf{\Sigma_q}\mathbf{G}_{qk}^H)) - \ln|\mathbf{S}_{q,k}| - N_{e,k} - \ln|\mathbf{M}_{e,q,k}|\,. \tag{10c}$$

Problem (9) is nonconvex with respect to (w.r.t) $(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K})$. However, it is easy to verify that problem (9) is convex w.r.t either $(\mathbf{\Sigma}_q, \mathbf{W}_q)$ or $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ (by checking its Hessian). A stationary point to problem (7) that satisfies its K.K.T optimality conditions then can be found by solving (9) sequentially w.r.t $(\mathbf{\Sigma}_q, \mathbf{W}_q)$ and $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ [5, Section IV-B]. Specifically, in one iteration, problem (9) is solved w.r.t only $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ to find an optimal solution $\{\mathbf{S}_{q,k}^*\}_{k=0}^{K}$. Next, with $\{\mathbf{S}_{q,k}^*\}_{k=0}^{K}$ plugged in (10a), the problem in (9) is optimized w.r.t $(\mathbf{\Sigma}_q, \mathbf{W}_q)$ to find an optimal solution $(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*)$. This Alternating Optimization (AO) cycle continues until reaching a convergence point. The nth iteration of AO, i.e., $(\mathbf{\Sigma}_q^n, \mathbf{W}_q^n, \{\mathbf{S}_{q,k}^n\}_{k=0}^{K})$, is as follows:

$$(\mathbf{\Sigma_q^n}, \mathbf{W_q^n}) = \arg\max_{(\mathbf{\Sigma}_q, \mathbf{W}_q)\in\mathcal{F}_q} f_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}^{n-1}\}_{k=0}^{K}) \tag{11a}$$

$$\mathbf{S}_{q,0}^n \triangleq \arg\max_{\mathbf{S}_{q,0}\succeq 0} \varphi_q(\mathbf{\Sigma}_q^n, \mathbf{W}_q^n, \mathbf{S}_{q,0}) = (\mathbf{M}_q^n)^{-1} = \left(\mathbf{I} + \mathbf{H}_{qq}\mathbf{W}_q^n\mathbf{H}_{qq}^H + \sum_{\substack{r=1\\ r\neq q}}^{Q} \mathbf{H}_{rq}\left(\mathbf{\Sigma}_r^0 + \mathbf{W}_r^0\right)\mathbf{H}_{rq}^H\right)^{-1} \tag{11b}$$

$$\mathbf{S}_{q,k}^n \triangleq \arg\max_{\mathbf{S}_{q,k}\succeq 0} \varphi_{e,q,k}(\mathbf{\Sigma}_q^n, \mathbf{W}_q^n, \mathbf{S}_{q,k}) = \left(\mathbf{M}_{e,q,k}^n + \mathbf{G}_{qk}{\mathbf{\Sigma_q}}^n\mathbf{G}_{qk}^H\right)^{-1}$$
$$= \left(\mathbf{I} + \mathbf{G}_{qk}(\mathbf{\Sigma}_q^n + W_q^n)\mathbf{G}_{qk}^H + \sum_{\substack{r=1\\ r\neq q}}^{Q} \mathbf{G}_{rk}\left(\mathbf{\Sigma}_r^0 + \mathbf{W}_r^0\right)\mathbf{G}_{rk}^H\right)^{-1},\ k \neq 0 \tag{11c}$$

where $\mathbf{\Sigma}_r^0$ and $\mathbf{W}_r^0$ (for $r \neq q$) denote the received interference components at the qth receiver prior to solving (9). Incorporating (11b) and (11c) in (11a), the solution to the convex problem (11a) can be found using a convex optimization solver. Notice that in (11b) and (11c), the users do not coordinate with each other in the middle of finding a stationary point for (9), for all $q \in \mathbb{Q}$. Hence, the terms $\mathbf{\Sigma}_r^0$ and $\mathbf{W}_r^0$, $r \neq q$ remain constant during the AO iterations. To solve problem (9) faster, the authors in [5] solved the smooth approximation of (7) based on the log-sum-exp inequality [16, chapter 3.1.5], which states that

$$\max\{a_1, \ldots, a_K\} \leq \frac{1}{\beta}\ln(\sum_{k=1}^{K} e^{\beta a_k}) \leq \max\{a_1, \ldots, a_K\} + \frac{1}{\beta}\ln K. \tag{12}$$

where $a_k \in \mathbb{R}$ and $\beta > 0$. Applying (12) to (4), we can write problem (7) as

$$\begin{aligned} \underset{\mathbf{\Sigma}_q, \mathbf{W}_q}{\text{maximize}}\ & \bar{R}_{s,q}(\mathbf{\Sigma}_q, \mathbf{W}_q) \\ \text{s.t.}\quad & (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q,\ q \in \mathbb{Q} \end{aligned} \tag{13}$$

where

$$\bar{R}_{s,q}(\mathbf{\Sigma}_q, \mathbf{W}_q) \triangleq C_q(\mathbf{\Sigma}_q, \mathbf{W}_q) - \frac{1}{\beta}\ln\Big(\sum_{k=1}^{K}\exp\left\{\beta C_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q)\right\}\Big),\ q \in \mathbb{Q}. \tag{14}$$

Hence, we can do the same reformulation procedure taken in (9) to end up with the following smooth reformulation [5]:

$$\begin{aligned} &\underset{\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_q}{\text{maximize}}\ \bar{f}_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K}), \\ &\text{s.t.} \qquad (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q,\ \mathbf{S}_k \succeq 0, q \in \mathbb{Q}, k \in \mathbb{K} \end{aligned} \tag{15}$$

where

$$\bar{f}_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K}) \triangleq \varphi_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,0}) - \frac{1}{\beta}\ln\Big(\sum_{k=1}^{K} e^{\beta\varphi_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,k})}\Big). \tag{16}$$

with φ_q and $\varphi_{e,q,k}$ defined in (10b) and (10c), respectively. Hence, the AO iteration in (11a) changes to

$$(\mathbf{\Sigma}_\mathbf{q}^\mathbf{n}, \mathbf{W}_\mathbf{q}^\mathbf{n}) = \arg\max_{(\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q} \bar{f}_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \left\{\mathbf{S}_{q,k}^{n-1}\right\}_{k=0}^{K}), \tag{17}$$

while $\left\{\mathbf{S}_{q,k}^{n-1}\right\}_{k=0}^{K}$ remain the same as (11b) and (11c)[4]. After plugging (11b) and (11c) into (17), the solution to (17) at the nth iteration is computed using the Projected Gradient (PG) algorithm. The lth iteration of PG algorithm while solving (17) is as follows.

$$\begin{pmatrix} \hat{\mathbf{\Sigma}}_q^{n,l+1} \\ \hat{\mathbf{W}}_q^{n,l+1} \end{pmatrix} = \mathtt{Proj}_{\mathcal{F}_q}\begin{pmatrix} \mathbf{\Sigma}_q^{n,l} + \alpha_l \nabla_{\mathbf{\Sigma}_q} \bar{f}_q^{n,l} \\ \mathbf{W}_q^{n,l} + \alpha_l \nabla_{\mathbf{W}_q} \bar{f}_q^{n,l} \end{pmatrix}, \tag{18}$$

$$\begin{pmatrix} \mathbf{\Sigma}_q^{n,l+1} \\ \mathbf{W}_q^{n,l+1} \end{pmatrix} = \begin{pmatrix} \mathbf{\Sigma}_q^{n,l} \\ \mathbf{W}_q^{n,l} \end{pmatrix} + \varepsilon_l \begin{pmatrix} \hat{\mathbf{\Sigma}}_q^{n,l+1} - \mathbf{\Sigma}_q^{n,l} \\ \hat{\mathbf{W}}_q^{n,l+1} - \mathbf{W}_q^{n,l} \end{pmatrix}, \tag{19}$$

where α_l and ε_l are step sizes that can be determined using Wolfe conditions for PG method [17]; $\mathtt{Proj}_{\mathcal{F}_q}$ is the projection operator to the set $\mathcal{F}_q$, which can be written as

$$\mathtt{Proj}_{\mathcal{F}_q}\begin{pmatrix} \tilde{\mathbf{\Sigma}} \\ \tilde{\mathbf{W}} \end{pmatrix} = \min_{\mathbf{W}, \mathbf{\Sigma} \in \mathcal{F}_q} ||\mathbf{W} - \tilde{\mathbf{W}}||_F^2 + ||\mathbf{\Sigma} - \tilde{\mathbf{\Sigma}}||_F^2; \tag{20}$$

and $(\nabla_{\mathbf{\Sigma}_q} \bar{f}_q^{n,l}, \nabla_{\mathbf{W}_q} \bar{f}_q^{n,l}) = \Big(\nabla_{\mathbf{\Sigma}_q} \bar{f}_q(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \left\{\mathbf{S}_{q,k}^{n-1}\right\}_{k=0}^{K}),$

[4]As far as optimality is concerned, it is shown in [5] that in the single-user scenario, the limit point of AO iterations done using (17), (11b), and (11c) are very close to the solutions found from AO iterations done using (11a), (11b), and (11c).

$\nabla_{\mathbf{W}_q} \bar{f}_q(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \{\mathbf{S}_{q,k}^{n-1}\}_{k=0}^K)\Big)$ where

$$\nabla_{\mathbf{\Sigma}_q} \bar{f}_q(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \{\mathbf{S}_{q,k}^{n-1}\}_{k=0}^K) =$$

$$\mathbf{H}_{qq}^H(\mathbf{M}_q^{n,l} + \mathbf{H}_{qq}\mathbf{\Sigma}_q^{n,l}\mathbf{H}_{qq}^H)^{-1}\mathbf{H}_{qq} - \sum_{k=1}^{K} \rho_{q,k}^{n,l}\mathbf{G}_{q,k}^H\mathbf{S}_{q,k}^{n-1}\mathbf{G}_{q,k}, \tag{21a}$$

$$\mathbf{M}_q^{n,l} = \mathbf{I} + \mathbf{H}_{qq}\mathbf{W}_q^{n,l}\mathbf{H}_{qq}^H + \sum_{\substack{r=1 \\ r\neq q}}^{Q} \mathbf{H}_{rq}\left(\mathbf{\Sigma}_r^0 + \mathbf{W}_r^0\right)\mathbf{H}_{rq}^H, \tag{21b}$$

$$\rho_{q,k}^{n,l} = \frac{e^{\beta\varphi_{e,q,k}(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \mathbf{S}_{q,k}^{n-1})}}{\sum_{j=1}^{K} e^{\beta\varphi_{e,q,j}(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \mathbf{S}_{q,j}^{n-1})}}, \tag{21c}$$

$$\nabla_{\mathbf{W}_q} \bar{f}_q(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \{\mathbf{S}_{q,k}^{n-1}\}_{k=0}^K) =$$
$$\mathbf{H}_{qq}^H\Big((\mathbf{M}_q^{n,l} + \mathbf{H}_{qq}\mathbf{\Sigma}_q^{n,l}\mathbf{H}_{qq})^{-1} - \mathbf{S}_{q,0}^{n-1}\Big)\mathbf{H}_{qq} +$$
$$\sum_{k=1}^{K} \rho_{q,k}^{n,l}\mathbf{G}_{qk}^H\left((\mathbf{M}_{e,q,k}^{n,l})^{-1} - \mathbf{S}_{q,k}^{n-1}\right)\mathbf{G}_{qk}, \tag{21d}$$

$$\mathbf{M}_{e,q,k}^{n,l} = \mathbf{I} + \mathbf{G}_{qk}\mathbf{W}_q^{n,l}\mathbf{G}_{qk}^H + \sum_{\substack{r=1 \\ r\neq q}}^{Q} \mathbf{G}_{rk}\left(\mathbf{\Sigma}_r^0 + \mathbf{W}_r^0\right)\mathbf{G}_{rk}^H. \tag{21e}$$

The projection in (20) can be efficiently computed according to [5, Fact 1]. We refer to the game where the actions of the players are defined by (51) as the proposed smooth game. Now that we have the response of each user, we can analyze the dynamics of the proposed smooth game.

A pseudo-code of the proposed smooth game mentioned so far is shown in Algorithm 1. As mentioned earlier, finding a stationary point for (51) for each user consists of two nested loops. The inner loop involves the gradient projection which is shown in (18) and (19) (i.e., the loop in Line 6 of Algorithm 1). Once the optimal solution to inner loop is found, one AO iteration is done by recalculating $\{\mathbf{S}_{q,k}\}_{k=0}^K$ according to (11b) and (11c) in the outer loop (i.e., Line 4). After the AO iterations converge to a stationary point, the users begin their transmissions using the computed precoders of information signal and AN[5]. Therefore, one round of this competitive secrecy rate maximization is done. Notice that according to Line 2, the players will be notified of actions of each other (i.e., recalculate the received interference) only after the AO iterations has converged[6]. The last round of the game will be the one where the convergence is reached.

[5]Although the optimization of covariance matrices of information signal and AN has been taken into account so far, the precoders can be found using eigenvalue decomposition.

[6]Such procedure in Line 2 of Algorithm 1 also explains the reason why $\mathbf{W}_r^0$ and $\mathbf{\Sigma}_r^0$ in (11) and (21) remain constant during AO iterations.

Algorithm 1 Proposed Smooth Game

Initialize: $\mathbf{\Sigma}_q^{1,1}$, $\mathbf{W}_q^{1,1}$, $\text{tr}(\mathbf{\Sigma}_q^{1,1}+\mathbf{W}_q^{1,1}) < P_q$, $\forall q \in \mathbb{Q}$

repeat
Each link q computes $\mathbf{M}_q$, $\mathbf{M}_{e,q,k}$, $\forall k \in \mathbb{K}$ locally
  for q =1,...,Q **do**
    for n = 1,... **do**
      Compute $\mathbf{S}_{q,k}^{n-1}$, $k = 0, \ldots, K$
        for l = 1,... **do**
          Compute $\varphi_{e,q,k}(\mathbf{\Sigma}_q^{n,l}, \mathbf{W}_q^{n,l}, \mathbf{S}_{q,k}^{n-1})$, $\mathbf{M}_q^{n,l}$, $\mathbf{M}_{e,q,k}^{n,l}$, $\forall (q,k)$
          Compute $(\mathbf{\Sigma}_q^{n,l+1}, \mathbf{W}_q^{n,l+1})$ using (18)-(21) % Use Wolfe conditions
        end for
    end for
  end for
until Convergence to QNE

IV. Existence and Uniqueness of the QNE

Before we begin to analyze the existence and uniqueness of the QNE, we review fundamentals of variational inequality theory as the basis of our analyses.

Variational Inequality Theory: Let $F : \mathcal{Q} \to \mathbb{R}^N$ be a vector-valued continuous real function, where $N > 1$ and $\mathcal{Q} \subseteq \mathbb{R}^N$ is a nonempty, closed, and convex set. The variational inequality $\text{VI}(F, \mathcal{Q})$ is the problem of finding a vector x^* such that

$$(x - x^*)^T F(x^*) \geq 0, \quad \forall x \in \mathcal{Q}. \tag{22}$$

The relation between variational inequality and game theory is summarized in the following theorem:

Theorem 1. *[11, Chapter 2] Consider Q players in a noncooperative game with utility function $f_q(x)$ for the qth player (not to be confused with the f_q defined in (9)), where $x \in \mathcal{Q}$ and $x = [x_1, x_2, ..., x_Q]^T$, x_q is the qth player's strategy, and $f_q(x)$ is concave w.r.t x_q for all q. The set $\mathcal{Q}$ is comprised of all strategy sets (i.e., $\mathcal{Q} = \prod_{q=1}^{Q} \mathcal{Q}_q$, where $\mathcal{Q}_q$ is the qth player's strategy set). Assuming the differentiability of $f_q(x)$ w.r.t x_q and that $\mathcal{Q}_q$ is a closed and convex set for all q, the vector x^* is the NE of the game if for $F(x) = [-\nabla_{x_1} f_1(x), -\nabla_{x_2} f_2(x), ..., -\nabla_{x_Q} f_Q(x)]^T$ we have:*

$$(x - x^*)^T F(x^*) \geq 0, \quad \forall x \in \mathcal{Q}.$$

□

A. Variational Inequality in Complex Domain

The theory of VI mentioned in (22) assumes that $\mathcal{Q} \subseteq \mathbb{R}^n$. However, this assumption might not be of our interest because the strategies of the players in our proposed game are two complex matrices (i.e., $\mathbf{\Sigma}_q$ and $\mathbf{W}_q$). Therefore, an alternative definition for VI in complex domain is needed. We use the definitions derived by the authors in [18] to define VI in complex domain.

Minimum Principle in Complex Domain: Consider the following optimization

$$\begin{aligned} &\underset{\mathbf{Z}}{\text{minimize}} \quad f(\mathbf{Z}) \\ &\text{s.t.} \quad \mathbf{Z} \in \mathcal{K} \end{aligned} \tag{23}$$

where $f : \mathcal{K} \to \mathbb{R}$ is convex and continuously differentiable on $\mathcal{K}$ where $\mathcal{K} \subseteq \mathbb{C}^{N' \times N}$, $N' > 1$, and $N > 1$. $X \in \mathcal{K}$ is an optimal solution to (23) if and only if we have [18, lemma23]

$$\langle \mathbf{Z} - \mathbf{X}, \nabla_{\mathbf{Z}} f(\mathbf{X}) \rangle \geq 0, \ \forall \mathbf{Z} \in \mathcal{K}. \tag{24}$$

where $\langle \mathbf{A}, \mathbf{B} \rangle = \text{Re}\left(Tr\left(\mathbf{A}^H \mathbf{B}\right)\right)$.

1) VI in Complex Domain: Using the definition of minimum principle in complex domain, we can now define the VI problem in the domain of complex matrices. For a complex-valued matrix $F^{\mathbb{C}}(\mathbf{Z}) : \mathcal{K} \to \mathbb{C}^{N' \times N}$ where $\mathcal{K} \subseteq \mathbb{C}^{N' \times N}$, the VI in the complex domain is the problem of finding a complex matrix $\mathbf{Y}$ such that the following is satisfied [18, Definition 25]

$$\left\langle \mathbf{Z} - \mathbf{Y}, F^{\mathbb{C}}(\mathbf{Y}) \right\rangle \geq 0, \ \forall \mathbf{Z} \in \mathcal{K}. \tag{25}$$

B. Quasi-Nash Equilibrium

It should be emphasized that the optimization problem of each player mentioned in (13) is nonconvex. Hence, the solution found for each link by solving (51) at Line 10 of Algorithm 1 is only a stationary point of problem (13). As a consequence, the traditional concepts of concave games used in proving the existence of a NE are not applicable here. Specifically, according to [9], the quasi-concavity of each player's utility w.r.t his strategy is required in proving the existence of a NE; an assumption that is not true in our game. Instead, we analyze the proposed (nonconvex) smooth game based on the relaxed equilibrium concept of QNE [10]. In the following, a formal definition of QNE is given [10].

Consider a noncooperative game with Q player each of whose strategies are restricted by some private constraints denoted as

$$\mathcal{X}_q = \{x_q \in X_q | h_q(x_q) \leq 0\}. \tag{26}$$

The set X_q is a convex set, and $h_q : \xi_q \to \mathbb{R}^{l_q}$ is a continuously differentiable mapping on the open convex set ξ_q containing X_q. No convexity assumption is made on h_q. Hence, although X_q is a convex set, $\mathcal{X}_q$ is not necessarily so. Player q has an objective function $g_q : \xi \to \mathbb{R}$, assumed to be continuously differentiable where $\xi = \prod_{q=1}^{Q} \xi_q$. The action of each player is formulated as follows:

$$\begin{aligned} &\underset{x_q \in \mathcal{X}_q}{\text{minimize}} \ g_q(x_q, x_{-q}) \\ &\text{s.t. } x_q \in \mathcal{X}_q. \end{aligned} \tag{27}$$

Obviously, the equivalent formulation can be written for when the action of each player is maximizing an objective (e.g., utility). Given the actions of other players, i.e., x^*_{-q}, and provided that a constraint qualification (CQ) condition holds at a point x^*_q, a necessary condition for x^*_q to be an optimal point of player q's optimization problem (i.e., action) is the existence of a nonnegative constant vector $\mu^*_q \in \mathbb{R}^{l_q}_+$ such that

$$\nabla_{x_q} L_q(x^*_q, x^*_{-q}, \mu^*_q) = \nabla_{x_q} g_q(x^*_q, x^*_{-q}) + \mu^{*T}_q \nabla_{x_q} h_q(x^*_q) = 0, \tag{28a}$$

$$\mu^{*T}_q h_q(x^*_q) = 0, \tag{28b}$$

$$h_q(x^*_q) \leq 0, \ x_q \in X_q. \tag{28c}$$

If any CQ is satisfied at x^*_q, the optimality conditions in (28) can be written as a VI over the set X_q; that is, the necessary condition for x^*_q to be an optimal solution to player q's optimization problem is if x^*_q solves VI$(\nabla_{x_q} L_q(\bullet, x^*_{-q}, \mu^*_q), X_q)$ [11, Proposition 1.3.4]. Furthermore, the existence of a nonnegative vector μ^*_q together with the complementarity of μ^*_q and $h_q(x^*_q)$ can be interpreted as μ^*_q being such that

$$-(\mu_q - \mu^*_q)^T h_q(x^*_q) \geq 0, \ \forall \mu_q \in \mathbb{R}^{l_q}_+. \tag{29}$$

Clearly, if $h_q(x^*_q)$ is not binding, i.e., $h_q(x^*_q) < 0$, then $\mu^*_q = 0$ satisfies (29). Furthermore, when $h_q(x^*_q)$ is binding, i.e., $h_q(x^*_q) = 0$, inequality (29) is trivially satisfied for all $\mu_q \in \mathbb{R}^{l_q}_+$. Hence, using (29) and the fact that x^*_q solves VI$(\nabla_{x_q} L_q(\bullet, x^*_{-q}, \mu^*_q), X_q)$, the pair (x^*_q, μ^*_q) solves the following VI:

$$\begin{pmatrix} x_q - x^*_q \\ \mu_q - \mu^*_q \end{pmatrix}^T \Gamma_q(x, \mu_q) \geq 0, \ \forall (x_q, \mu_q) \in \mathcal{R}_q = X_q \times \mathbb{R}^{l_q}_+ \tag{30}$$

where

$$\Gamma_q(x,\mu_q)=\begin{pmatrix}\nabla_{x_q}L_q(\bullet,x^*_{-q},\mu^*_q)\\ -h_q(x^*_q)\end{pmatrix}. \tag{31}$$

Notice that although it might seem that VI($\nabla_{x_q}L_q(\bullet,x^*_{-q},\mu^*_q),X_q$) and (29) cannot be combined to build (30), using the fact that VI is a generalized definition of a set-valued mapping[7], we are able to justify (30). it can be proved that for the set-valued mappings $N_{X_q}(x_q)$ and $N_{\mathbb{R}^{l_q}_+}(\mu_q)$, we have $N_{X_q\times\mathbb{R}^{l_q}_+}(x_q,\mu_q)=N_{X_q}(x_q)\times N_{\mathbb{R}^{l_q}_+}(\mu_q)$ [19]. The same conclusion holds for VI problems. Hence, inequality (30) can be deduced.

Concatenating the inequality in (30) over the set of players, the QNE can be defined as follows:

Definition 1. *The QNE is the pair* $\left(x^*_q,\mu^*_q\right)$, $q=1\ldots,Q$, *that satisfies the following inequality:*

$$\left(\begin{pmatrix}x_q-x^*_q\\ \mu_q-\mu^*_q\end{pmatrix}_{q=1}^{Q}\right)^T(\Gamma_q(x,\mu_q))_{q=1}^{Q}\geq 0,$$
$$\forall(x_q,\mu_q)_{q=1}^{Q}\in\prod_{q=1}^{Q}\mathcal{R}_q=\prod_{q=1}^{Q}(X_q\times\mathbb{R}^{l_q}_+) \tag{32}$$

where $(\bullet)_{q=1}^{Q}$ *denotes a column vector.*

Notice that the set $\prod_{q=1}^{Q}\mathcal{R}_q$ is a convex set, and if the actions of each player is a convex program, the QNE reduces to NE. In our scenario, since the private constraints for each player is a convex set, we embedded the private constraints into the set $\mathcal{R}_q$ defined in (32). We need to emphasize the fact that the constant vectors μ^*_q for all q can only be defined if the optimization problem of each player satisfies some CQ conditions. For players with convex problems, these constant vectors are trivially satisfied since the K.K.T conditions are necessary and sufficient conditions of optimality in convex programs.

One intuition that can be given on the concept of QNE is as follows. QNE is point where no player has an incentive to unilaterally change his strategy because any change makes a player not satisfy the K.K.T conditions of his problem. This is in contrast with the definition of NE in which the lack of incentives at NE is because of losing optimality. Again, optimality and satisfying the K.K.T conditions are equivalent when players solve convex programs.

C. Analysis of QNE

According to the aforementioned definition, the QNEs are tuples that satisfy the K.K.T conditions of all players' optimization problems. Under a constraint qualification, stationary points of each player's optimization problem satisfy its K.K.T conditions. To begin the analysis of the QNE, we first show that the stationary point found using AO mentioned previously (i.e., Line 4-10 of Algorithm 1) satisfies the K.K.T conditions of (13).

Proposition 1. *For the qth link,* $q\in\mathbb{Q}$*, the stationary point found using AO (i.e., Line 4-10 of Algorithm 1) satisfies the K.K.T conditions of* (13)*.*

Proof: See Appendix A. ∎

Now that the K.K.T optimality of the stationary point found by AO iterations is proved, we rewrite the K.K.T conditions of all players to a proper VI problem [10]. The solution(s) to the obtained VI is the QNE(s) of the proposed smooth game. For the proposed smooth game defined using (51), we can

[7]A point-to-set map, also called a multifunction or a set-valued map, is a map N from $\mathbb{R}^n$ into the power set of $\mathbb{R}^n$, i.e., for every $x\in\mathbb{R}^n$, $N_{\mathbb{R}^n}(x)$ is a (possibly empty) subset of $\mathbb{R}^n$ [11, Chapter 2.1.3]. To avoid confusion, note that the definition of a set-valued map is fundamentally different from that of a vector function such as $h_q(x_q)$ defined in (26).

establish the following VI to characterize the QNE points. Let the QNE point be as follows

$$\mathbf{Y} = \{\mathbf{Y}_q\}_{q=1}^{Q} \triangleq [\mathbf{\Sigma}^T, \mathbf{W}^T]^T = \{[\mathbf{\Sigma}_q^T, \mathbf{W}_q^T]^T\}_{q=1}^{Q} \tag{33}$$

where $\{[\mathbf{\Sigma}_q^T, \mathbf{W}_q^T]^T\}_{q=1}^{Q} = [\mathbf{\Sigma}_1^T, \mathbf{W}_1^T, \mathbf{\Sigma}_2^T, \mathbf{W}_2^T, \ldots, \mathbf{\Sigma}_Q^T, \mathbf{W}_Q^T,]^T$. The function $F^{\mathbb{C}}(Z)$ is written as

$$F^{\mathbb{C}} = F^{\mathbb{C}}(\mathbf{\Sigma}, \mathbf{W}, \mathbf{S}) = \left\{F_q^{\mathbb{C}}(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K})\right\}_{q=1}^{Q} \triangleq$$
$$\left\{\left[-(\nabla_{\mathbf{\Sigma}_q}\bar{f}_q)^T, -(\nabla_{\mathbf{W}_q}\bar{f}_q)^T\right]^T\right\}_{q=1}^{Q} \tag{34}$$

where the terms $\nabla_{\mathbf{\Sigma}_q}\bar{f}_q$ and $\nabla_{\mathbf{W}_q}\bar{f}_q$ are given in (21). Therefore, the system of inequalities indicated as $VI(F^{\mathbb{C}}, \mathcal{K})$ can be established according to (25), where $\mathcal{K} = \prod_{q=1}^{Q}\mathcal{F}_q$. Furthermore, for a given response $\mathbf{\Sigma}_q$ and $\mathbf{W}_q$, the solutions of $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ are uniquely determined by (11b) and (11c) for all q. Hence, from now on, we assume that the values of $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ are already plugged into $F_q^{\mathbb{C}}\left(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K}\right)$, so we drop the term $\{\mathbf{S}_{q,k}\}_{k=0}^{K}$ in the subsequent equations for notational convenience.

In order to show that the K.K.T conditions are valid necessary conditions for a stationary solution of (13), an appropriate CQ must hold [20]. In this paper, we use the Slater's constraint qualification [20] as the strategy set of each player is a convex set. Moreover, at NE (if it exists) all of the players use their best responses, i.e., each player has found the optimal solution to his optimization problem and will not deviate from that. Since the optimal solution for each player also satisfies the K.K.T conditions, then NE must be a QNE [10]. In fact, the set of QNEs includes the NE.

D. Existence and Uniqueness of the QNE

To begin our analysis in this part, we consider the VI described by (25), (33), and (34) again. In the case of the domain of $\mathbf{Z}$ being square complex matrices, the definition of VI in complex domain can be further simplified to achieve the same form of VI in the real case (i.e., (22)). More specifically, let $F^{\mathbb{C}}$ be a $2N \times N$ matrix and let $\text{vec}(F^{\mathbb{C}}) \triangleq [(F_1)^T, \ldots, (F_N)^T]^T$ denote a $2N^2 \times 1$ vector where $F_i \triangleq [F^{\mathbb{C}}(\mathbf{Z})]_{:,i},\ i = 1, \ldots, N$, denotes the vector corresponding to the ith column of $F^{\mathbb{C}}(\mathbf{Z})$. Furthermore, let $\text{vec}(\mathbf{Z}) = [[\mathbf{Z}]_{:,1}^T, \ldots, [\mathbf{Z}]_{:,N}^T]^T$ be the vector version of the complex matrix $\mathbf{Z}$. Hence, the vector version of the VI in complex domain can be expressed as

$$(\text{vec}(\mathbf{Z}) - \text{vec}(\mathbf{Y}))^H \text{vec}(F^{\mathbb{C}}(\mathbf{Y})) \geq 0,\ \forall \mathbf{Z} \in \mathcal{K}. \tag{35}$$

In order to further simplify the VI in complex domain to be completely identical to the real case, we define $F^{\mathbb{R}} \triangleq [\text{Re}\left\{\text{vec}(F^{\mathbb{C}})\right\}^T, \text{Im}\left\{\text{vec}(F^{\mathbb{C}})\right\}^T]^T$ and $\mathbf{Z}^{\mathbb{R}} \triangleq [\text{Re}\left\{\text{vec}(\mathbf{Z})\right\}^T, \text{Im}\left\{\text{vec}(\mathbf{Z})\right\}^T]^T$ where Re{...} and $I\{...\}$ are the real and imaginary parts, respectively. Therefore, the real-vectorized representation of (25) can be written as

$$\left(\mathbf{Z}^{\mathbb{R}} - \mathbf{Y}^{\mathbb{R}}\right)^T \left(F^{\mathbb{R}}(\mathbf{Y}^{\mathbb{R}})\right) \geq 0,\ \forall \mathbf{Z}^{\mathbb{R}} \in \mathcal{K}^{\mathbb{R}},\ \text{where } \mathcal{K}^{\mathbb{R}} \subseteq \mathbb{R}^{2N^2}. \tag{36}$$

The vector form of (33) and (34) are as follows:

$$\text{vec}(\mathbf{Z}) = [\text{vec}(\bar{\mathbf{\Sigma}})^T, \text{vec}(\bar{\mathbf{W}})^T]^T = \left\{[\text{vec}(\bar{\mathbf{\Sigma}}_q)^T, \text{vec}(\bar{\mathbf{W}}_q)^T]^T\right\}_{q=1}^{Q} \tag{37}$$

$$\text{vec}(F^{\mathbb{C}}(\mathbf{Z}))=\left\{\left[\text{vec}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q)^T, \text{vec}(-\nabla_{\mathbf{W}_q}\bar{f}_q)^T\right]^T\right\}_{q=1}^{Q}. \tag{38}$$

Hence, the vector form of the complex VI problem $VI(F^{\mathbb{C}}, \mathcal{K})$ can be written as

$$\left([\text{vec}(\mathbf{\Sigma})^T, \text{vec}(\mathbf{W})^T]^T - [\text{vec}(\bar{\mathbf{\Sigma}})^T, \text{vec}(\bar{\mathbf{W}})^T]^T\right)^H \text{vec}(F^{\mathbb{C}}(\bar{\mathbf{\Sigma}}, \bar{\mathbf{W}})) \geq 0. \tag{39}$$

$$\left(\left[\mathbf{\Sigma}^{\mathbb{R}^T}, \mathbf{W}^{\mathbb{R}^T}\right] - \left[\bar{\mathbf{\Sigma}^{\mathbb{R}}}^T, \bar{\mathbf{W}^{\mathbb{R}}}^T\right]\right) F^{\mathbb{R}} \geq 0,\ \forall (\mathbf{\Sigma}^{\mathbb{R}}, \mathbf{W}^{\mathbb{R}}) \in \mathcal{K}^{\mathbb{R}},\ \mathcal{K}^{\mathbb{R}} \subseteq \mathbb{R}^m. \tag{40}$$

Hence, the equivalent real-vectorized representation of the VI in (25) that complies with the definition in (22) can be determined as (40) where $m \triangleq \sum_{q=1}^{Q} 2N_{T_q}^2$. Note that the set of matrices $(\mathbf{\Sigma}_1, \ldots, \mathbf{\Sigma}_Q, \mathbf{W}_1, \ldots, \mathbf{W}_Q)$ that are in $\mathcal{K} = \prod_{q=1}^{Q} \mathcal{F}_q$ are the ones whose real-vectorized versions will be inside $\mathcal{K}^{\mathbb{R}}$. Now that the proposed smooth game is modeled as a real-vectorized VI, we can use the following theorem to prove the existence of the QNE.

Theorem 2. *The proposed smooth game, where the actions of each player is given by* (51) *admits at least one QNE.*

Proof: See Appendix B. ∎

The uniqueness of the QNE is discussed in the following theorem:

Theorem 3. *The proposed smooth game characterized by* (51) *has a unique QNE if*

$$\lambda_{q,\min} > \sum_{\substack{q=1 \\ q \neq l}}^{Q} |||D_{Z_l} F_q^{\mathbb{C}}(Z_q)|||_2, \;\; q \in \mathbb{Q} \tag{41}$$

where $\lambda_{q,\min}$ *is the smallest eigenvalue of* $D_{Z_q} F_q^{\mathbb{C}}(Z_q)$*, and* $D_{Z_l} F_q^{\mathbb{C}}(Z_q) \triangleq \frac{\partial \; vec\left(F_q^{\mathbb{C}}(Z_q)\right)}{\partial \; vec(Z_l)^T}$*, for all* $q, l \in \mathbb{Q}^2$*, is defined as*

$$D_{Z_l} F_q^{\mathbb{C}}(Z_q) \triangleq \begin{bmatrix} D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{\Sigma}_q} \bar{f}_q) & D_{\mathbf{W}_l}(-\nabla_{\mathbf{\Sigma}_q} \bar{f}_q) \\ D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{W}_q} \bar{f}_q) & D_{\mathbf{W}_l}(-\nabla_{\mathbf{W}_q} \bar{f}_q) \end{bmatrix}. \tag{42}$$

Proof: See Appendix C. ∎

V. Analysis of the Proposed Game in the Presence of Multiple QNEs

A. On the Convergence of Algorithm 1

The conditions for the uniqueness of the QNE do not guarantee the convergence of Algorithm 1 to a (unique) QNE. Since the optimization of each player is nonconvex, only stationary points of players' utilities could be achieved. Hence, solving each player's optimization problem using AO does not necessarily lead to the best response of each player. This hinders us from proving the convergence of Algorithm 1. However, we verified the convergence via simulations. In this section, we present a slightly modified algorithm, namely, the gradient-response algorithm with proof of convergence. Furthermore, the gradient-response algorithm paves the way for further performance improvements introduced later in this paper.

B. Gradient-Response Algorithm

A solution to the VI in (40) can be characterized by the following iteration [11, Chapter 12]:

$$x^{(i+1)} = \Pi_{\mathcal{K}^{\mathbb{R}}} \left(x^{(i)} - \gamma F^{\mathbb{R}}(x^{(i)}, \{S_{q,k}^{(i)}\}_{k=0}^{K}) \right) \tag{43}$$

where $\Pi_{\mathcal{K}^{\mathbb{R}}}$ is the projection to set $\mathcal{K}^{\mathbb{R}}$, $x = \left[\mathbf{\Sigma}^{\mathbb{R}^T}, \mathbf{W}^{\mathbb{R}^T} \right]^T$, the superscript (i) is the number of iterations, and $\gamma = \text{diag}([\gamma_1, \ldots, \gamma_m]^T)$ is a diagonal matrix which indicates the step size that each player takes in

the improving direction of his utility function. The solutions to $\{\mathbf{S}_{q,k}^{(i)}\}_{k=0}^{K}$ are as follows:

$$\mathbf{S}_{q,0}^{(i)} \triangleq (\mathbf{M}_q^{(i)})^{-1} = \left(\mathbf{I} + \mathbf{H}_{qq}\mathbf{W}_q^{(i)}\mathbf{H}_{qq}^H + \sum_{\substack{r=1 \\ r\neq q}}^{Q} \mathbf{H}_{rq}\left(\mathbf{\Sigma}_r^{(i-1)} + \mathbf{W}_r^{(i-1)}\right)\mathbf{H}_{rq}^H\right)^{-1}, \tag{44a}$$

$$\mathbf{S}_{q,k\neq 0}^{(i)} \triangleq \left(\mathbf{M}_{e,q,k}^{(i)} + \mathbf{G}_{qk}\mathbf{\Sigma_q}^{(i)}\mathbf{G}_{qk}^H\right)^{-1} = \left(\mathbf{I} + \mathbf{G}_{qk}(\mathbf{\Sigma}_q^{(i)} + W_q^{(i)})\mathbf{G}_{qk}^H + \sum_{\substack{r=1 \\ r\neq q}}^{Q} \mathbf{G}_{rk}\left(\mathbf{\Sigma}_r^{(i-1)} + \mathbf{W}_r^{(i-1)}\right)\mathbf{G}_{rk}^H\right)^{-1} \tag{44b}$$

where (44b) holds for $k \neq 0$. It is easy to confirm that the iteration in (43) is a simplified version of the projection done by each user in (18) and (19). Notice that the only difference of the gradient-response algorithm, characterized by iteration in (43), from Algorithm 1 is that at each round of the gradient-response algorithm, a player only does one iteration of the PG method (i.e., (18)) and one iteration according to (44). The real-vectorized version of the gradient-response algorithm is shown in (43). Since the values of $\{\mathbf{S}_{q,k}^{(i)}\}_{k=0}^{K}$ are uniquely determined for a given $x^{(i)}$, we drop the term $\{\mathbf{S}_{q,k}^{(i)}\}_{k=0}^{K}$ from the argument of $F^{\mathbb{R}}$ for notational convenience.

Assuming that $F^{\mathbb{R}}$ is *strongly monotone* (with modulus $c_s/2$)[8] and *Lipschitz continuous* (with constant L)[9] w.r.t $(\mathbf{\Sigma}_q, \mathbf{W}_q)$, the convergence to a unique solution follows if $\gamma_{i'} = d < \frac{c_s}{L^2},\ \forall i' = 1, \ldots, m$, where d is constant. Hence, the mapping $x \to \Pi_{\mathcal{K}^{\mathbb{R}}}\left(x - \gamma F^{\mathbb{R}}(x)\right)$ becomes a contraction mapping and the fixed points of this map are solutions of the VI in (40) [11, Chapter 12]. It turns out that sufficient conditions for the strong monotonicity of $VI(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ are in fact the same as the conditions derived in (41) for the uniqueness of the QNE[10]. Therefore, based on (43), a pseudo-code of the gradient-response algorithm is given in Algorithm 2. Note that the operation in Line 6 of Algorithm 2 is the same as the iteration in (43). In fact, since the set $\mathcal{K}^{\mathbb{R}}$ is a Cartesian product of players' strategies, the iteration in (43) can be easily converted back to its matrix form to have the following iteration:

$$\begin{pmatrix} \mathbf{\Sigma}_q^{(i+1)} \\ \mathbf{W}_q^{(i+1)} \end{pmatrix} = \mathtt{Proj}_{\mathcal{F}_q}\left(\begin{array}{c} \mathbf{\Sigma}_q^{(i)} + \gamma_q' \nabla_{\mathbf{\Sigma}_q} \bar{f}_q(\mathbf{\Sigma}_q^{(i)}, \mathbf{W}_q^{(i)}, \left\{\mathbf{S}_{q,k}^{(i)}\right\}_{k=0}^{K}) \\ \mathbf{W}_q^{(i)} + \gamma_q' \nabla_{\mathbf{W}_q} \bar{f}_q(\mathbf{\Sigma}_q^{(i)}, \mathbf{W}_q^{(i)}, \left\{\mathbf{S}_{q,k}^{(i)}\right\}_{k=0}^{K}) \end{array}\right), \forall q \in \mathbb{Q}. \tag{45}$$

Notice that γ_q' is a diagonal matrix that can obtained by dividing the matrix γ into Q block-diagonal matrices. That is, with a slight abuse of notations, $\gamma = \text{diag}([\gamma_1, \ldots, \gamma_m]^T) = \gamma' = \text{diag}(\gamma_1', \ldots, \gamma_Q')$, $Q < m$. Therefore, the gradient response in (43) can be shown as an iteration that is done in each link, independent of other links. This is essentially a distributed implementation. The gradient-response algorithm is given in Algorithm 2.

[8]The notion of strong monotonicity is a basic definition in the topic of VI (see Appendix C).

[9]It can be seen from (18) and (19) that the power constraint of each user makes the variations of $\nabla_{\mathbf{\Sigma}_q}\bar{f}_q$ and $\nabla_{\mathbf{W}_q}\bar{f}_q$ bounded for all $q \in \mathbb{Q}$. Hence, $F^{\mathbb{R}}$ is Lipschitz continuous on $\mathcal{K}^{\mathbb{R}}$.

[10]More explanation can be found in Appendix C.

Algorithm 2 Gradient-Response Algorithm

Initialize: $\mathbf{\Sigma}_q^{(1)}$, $\mathbf{W}_q^{(1)}$, $\text{tr}(\mathbf{\Sigma}_q^{(1)} + \mathbf{W}_q^{(1)}) < P_q$, $\forall q$

1: **repeat** % superscript (i) indicates the iterations starting from here

2: Compute $\mathbf{M}_q$, $\mathbf{M}_{e,q,k}$, $\forall (q,k) \in \mathbb{Q} \times \mathbb{K}$

3: Compute $\mathbf{S}_{q,k}^{(i)}$, $\forall (q,k) \in \mathbb{Q} \times \mathbb{K}$

4: Compute $\varphi_{e,q,k}(\mathbf{\Sigma}_q^{(i)}, \mathbf{W}_q^{(i)}, \mathbf{S}_{q,k}^{(i)})$, $\forall (q,k) \in \mathbb{Q} \times \mathbb{K}$

5: **for** q =1,...,Q **do**

6: Compute $(\mathbf{\Sigma}_q^{(i+1)}, \mathbf{W}_q^{(i+1)})$ using (45)

7: **end for**

8: **until** Convergence to QNE

The convergence point of Algorithm 2 is a QNE of the game where players' actions are defined by (51). Specifically, assume that for $i \to \infty$, the convergence point is denoted as $(\bar{\mathbf{\Sigma}}, \bar{\mathbf{W}})$. Hence, we have for all $q \in \mathbb{Q}$

$$\bar{\mathbf{S}}_{q,0} = \arg \max_{\mathbf{S}_{q,0} \succeq 0} \varphi_q(\bar{\mathbf{\Sigma}}_q, \bar{\mathbf{W}}_q, \mathbf{S}_{q,0}) \tag{46a}$$

$$\bar{\mathbf{S}}_{q,k} = \arg \max_{\mathbf{S}_{q,k} \succeq 0} \varphi_{e,q,k}(\bar{\mathbf{\Sigma}}_q, \bar{\mathbf{W}}_q, \mathbf{S}_{q,k}), \ k \neq 0. \tag{46b}$$

The solution of (46a) and (46b) is the same as (44a) and (44b) for $i \to \infty$. By plugging the solutions of (46a) and (46b) in $\nabla_{\mathbf{\Sigma}_q} \bar{f}_q(\bar{\mathbf{\Sigma}}_q, \bar{\mathbf{W}}_q, \{\mathbf{S}_{q,k}\}_{k=0}^K)$ and $\nabla_{\mathbf{W}_q} \bar{f}_q(\bar{\mathbf{\Sigma}}_q, \bar{\mathbf{W}}_q, \{\mathbf{S}_{q,k}\}_{k=0}^K)$, the convergence point of Algorithm 2 is a QNE of the proposed game. Overall, by using the gradient-response algorithm, the uniqueness of the QNE and $\gamma_{i'} = d < \frac{c_s}{L^2}, \ \forall i' = 1, \ldots, m$ directly suggest the convergence of the iteration in (43). Hence, a separate proof for the convergence of Algorithm 2 is not needed.

The iteration proposed in (43) has two major issues. First, the Lipschitz constant of $F^{\mathbb{R}}(x)$ has to be known. Apart from being difficult to derive, the knowledge of Lipschitz constant requires a centralized computation. Second, the strong monotonicity of $F^{\mathbb{R}}$ cannot be always guaranteed. In fact, the conditions derived in (41) are very dependent on the channel gains and network topology. Hence, in most typical network scenarios, the inequality in (41) cannot be satisfied. This means that in some situations, the game might have more than one QNE. Consequently, the convergence of Algorithm 2 is in jeopardy. However, on the condition that $F^{\mathbb{R}}$ is *monotone*[11], which is a weaker condition than strong monotonicity, the ability to choose between multiple QNEs is possible. This means that the users are able to select the QNE that satisfies a certain design criterion, thus guaranteeing convergence in the case of multiple QNEs. Moreover, depending on the design criterion, the performance of the resulting QNE in terms of the achieved secrecy sum-rate can be improved. To do this, we first review the regularization methods proposed for VIs.

C. Tikhonov Regularization

The general idea of regularization techniques is to modify the players' utility functions such that the VI becomes strongly monotone (and hence easily solvable by using Algorithm 2), and the limit point of a sequence of solutions for the modified VI converges to some solution of the original VI. In Tikhonov regularization, the process of regularizing $\text{VI}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ involves solving a sequence of VIs, where the following iteration is characterized for a given ϵ [11, chapter 12]:

$$x^{(i+1)} = \Pi_{\mathcal{K}^{\mathbb{R}}} \left(x^{(i)} - \gamma^T \left(F^{\mathbb{R}}(x^{(i)}) + \epsilon x^{(i)} \right) \right). \tag{47}$$

The solution to (47) when $i \to \infty$ is denoted as $x(\epsilon)$. Given that $F^{\mathbb{R}}$ is monotone, solving a sequence of (strongly monotone) $\text{VI}(F^{\mathbb{R}}(x) + \epsilon x, \mathcal{K}^{\mathbb{R}})$'s while $\epsilon \to 0$ has a limit point, (i.e., $\lim_{\epsilon \to 0} x(\epsilon)$ exists) and that limit point is equal to least-norm solution of the $\text{VI}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ [11, Theorem 12.2.3].

[11] See Appendix C to recall the difference between monotonicity and strong monotonicity.

D. QNE Selection using Tikhonov Regularization

Generalizing the applicability of Tikhonov regularization, we are more interested in converging to the QNE that is more beneficial to the users. In our approach to QNE selection, we define benefit as when the selected QNE satisfies a particular design criterion. Let the set of solutions of VI$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ be denoted as SOL$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$. We want to select the NE that minimizes a strongly convex[12] function $\Phi(x) : \mathcal{K}^{\mathbb{R}} \rightarrow \mathbb{R}$. In fact, the QNE selection satisfies the following design criterion[13]

$$\begin{aligned} \text{minimize} \quad & \Phi(x) \\ \text{s.t.} \quad & x \in \text{SOL}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}). \end{aligned} \tag{48}$$

The optimization in (48) is convex because the monotonicity of $F^{\mathbb{R}}$ suggests that $SOL(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ is a convex set [11, Chapter 2]. The unique point that solves problem (48), is the solution to VI$(\nabla\Phi(x), \text{SOL}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}))$. However, as there is no prior knowledge on SOL$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ (i.e., QNEs are not known), this optimization cannot be solved easily. To overcome this issue, we modify the function $F^{\mathbb{R}}$ in VI$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ to

$$F_{\epsilon}^{\mathbb{R}} \triangleq F^{\mathbb{R}} + \epsilon\nabla\Phi(x). \tag{49}$$

As the function $\Phi(x)$ is a strongly convex function, its derivative w.r.t x is strongly monotone. Assuming that $F^{\mathbb{R}}$ is monotone, then the function $F_{\epsilon}^{\mathbb{R}}$ is strongly monotone and the solution to VI$(F_{\epsilon}^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$, namely, $x(\epsilon)$, is unique for all values of $\epsilon > 0$ (i.e., convergence to a QNE can be guaranteed). The iteration used for QNE selection is written as

$$x^{(i+1)} = \Pi_{\mathcal{K}^{\mathbb{R}}}\left(x^{(i)} - \gamma^{(i)}\left(F^{\mathbb{R}}(x^{(i)}) + \epsilon^{(j)}\nabla\Phi(x) + \theta^{(i)}(x^{(i)} - x^{(i-1)})\right)\right). \tag{50}$$

The iteration in (50) is the same as (47) with the difference that the multiplier of ϵ in (47) is replaced by $\nabla\Phi(x)$. The following theorem shows the potential of using (50) for QNE selection:

Theorem 4. *[11, pp. 1128 and Theorem 12.2.5] Consider VI$(F_{\epsilon}^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ with $x(\epsilon)$ as its solution. Assume that $\mathcal{K}^{\mathbb{R}}$ is closed and convex, and SOL$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ is nonempty. The following claims hold:*

- *The assumption that $\mathcal{K}^{\mathbb{R}}$ is closed and convex together with the nonemptiness of SOL$(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ (i.e., the existence of the QNE, proved in Theorem 2) are necessary and sufficient for $x_{\infty} = \lim_{\epsilon \rightarrow 0} x(\epsilon)$ to exist.*
- *Assuming that $F^{\mathbb{R}}$ is monotone[14], x_{∞} is the solution of VI$(\nabla\Phi(x), SOL(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}))$. This means that a QNE among several QNEs can be selected[15].* □

E. Guaranteeing Monotoncity of $F^{\mathbb{R}}$ in Tikhonov Regularization

Theorem 4 requires $F^{\mathbb{R}}$ to be monotone to be applicable. However, the monotonicity of $F^{\mathbb{R}}$, as highlighted by Theorem 3, depends on many factors such as the channels between different nodes in the network, meaning that it is not possible to always guarantee the monotonicity of $F^{\mathbb{R}}$. In order to guarantee the monotonicity, we add a strongly concave term to the utility of each player. Let this term be $-\frac{\tau_q}{2}\left(||\mathbf{\Sigma}_q - Y_{\mathbf{\Sigma}_q}||_F^2 + ||\mathbf{W}_q - Y_{\mathbf{W}_q}||_F^2\right)$ where $||.||_F$ indicates the Frobenius norm. Hence, the utility of each player defined in (51) will change to

$$\begin{aligned} \underset{\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_q}{\text{maximize}} \; & \bar{f}_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}\}_{k=0}^{K}) - \frac{\tau_q}{2}\left(||\mathbf{\Sigma}_q - Y_{\mathbf{\Sigma}_q}||_F^2 + ||\mathbf{W}_q - Y_{\mathbf{W}_q}||_F^2\right), \\ \text{s.t.} \quad & (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q, \; \mathbf{S}_k \succeq 0, \; q \in \mathbb{Q}, \; k \in \mathbb{K} \end{aligned} \tag{51}$$

[12]A strongly convex function is a function whose derivative is strongly monotone. We use the definitions of [21] to distinguish between different types of convexity.

[13]The discussion on how we determine the function $\Phi(x)$ will be tackled in Section VI-B.

[14]Later on, we elaborate on the monotonicity assumption for $F^{\mathbb{R}}$ (cf. Section VIII-A).

[15]We emphasize that by QNE selection, the players are maximizing their (modified) utility functions. Hence, the noncooperative nature of the game is preserved.

where $Y_{\mathbf{\Sigma}_q}$ and $Y_{\mathbf{W}_q}$ are complex constants which will be explained later. With this modification on the utility of each player, a new VI problem, $VI(F_\tau^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ is established where:

$$F_\tau^{\mathbb{R}}(x) = F^{\mathbb{R}}(x) + \tau(x - y) \tag{52}$$

where y is the vector that contains the vectorized versions of $Y_{\mathbf{\Sigma}_q}$ and $Y_{\mathbf{W}_q}$, and $\tau = diag(\tau_1, \tau_2, \ldots, \tau_m)$ is an $m \times m$ diagonal matrix, and $F^{\mathbb{R}}$ is not a function of y. This perturbation is also known as Proximal Point regularization method [11, Chapter 12.3.2]. Recalling Definition 1, the augmented Jacobian matrix of $F_\tau^{\mathbb{R}}(x)$, namely as $\mathcal{J}_\tau$, is as follows

$$\mathcal{J}_\tau \triangleq \mathcal{J} + \tau I \tag{53}$$

where $\mathcal{J}$ is the augmented Jacobian matrix of $F^{\mathbb{R}}$ and I is the identity matrix. Considering the matrix τ as a free parameter, we can choose a suitable value for each diagonal element of τ, such that the matrix $\mathcal{J}_\tau$ becomes a diagonally dominant matrix. In the following we exploit the diagonal dominance of $\mathcal{J}_\tau$ to establish the monotonicity property of $F_\tau^{\mathbb{R}}$[16].

Let $D(d_i, [\mathcal{J}_\tau]_{ii}),\ i = 1, \ldots, m$ be the closed disc centered at $[\mathcal{J}_\tau]_{ii}$ with radius $d_i = \sum_{j \neq i} |[\mathcal{J}_\tau]_{ij}|$, where $[.]_{ii}$ denotes the diagonal element and $[\mathcal{J}_\tau]_{ii} = [\mathcal{J}]_{ii} + \tau_i$. Using the Gerschgorin circle theorem [22], for all $i = 1, \ldots, m$, every eigenvalue of $\mathcal{J}_\tau$ is within at least one of the discs. We also know that for the function $F_\tau^{\mathbb{R}}$, in order to be monotone, the matrix $\mathcal{J}_\tau$ has to be APSD (cf. Appendix C). Hence, provided that a suitable value for τ_i is chosen for all $i = 1, \ldots, m$, all the radii of the Gershgorin circles must be less than their respective diagonal elements, ensuring that $\mathcal{J}_\tau$ remains APSD. Using this fact, the value for τ_i that guarantees $\mathcal{J}_\tau$ to be APSD is

$$\tau_i \geq d_i - \mathcal{J}_{ii},\ \forall i. \tag{54}$$

Therefore, using the condition (54) with equality, the matrix $\mathcal{J}_\tau$ becomes an APSD matrix, and consequently, $F_\tau^{\mathbb{R}}$ becomes monotone. Therefore, the Tikhonov regularization changes to solving the problem $VI(F_{\tau,\epsilon}^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ where

$$F_{\tau,\epsilon}^{\mathbb{R}} \triangleq F^{\mathbb{R}}(x) + \tau(x - y) + \epsilon^{(j)} \nabla \Phi(x) \tag{55}$$

Building upon the perturbation in (52), we can now use $F_\tau^{\mathbb{R}}$ instead of $F^{\mathbb{R}}$ in the original VI in (40) which makes us able to use Tikhonov regularization and perform equilibrium selection. One might argue that using $F_\tau^{\mathbb{R}}$ instead of $F^{\mathbb{R}}$ is actually creating a new game with different solutions. In the following we give a property that makes the use of $F_\tau^{\mathbb{R}}$ reasonable. It can be easily seen that the perturbation $F_\tau^{\mathbb{R}}$ does not change the fact that the NE in $VI(F_\tau^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ still exists, i.e., the set $\text{SOL}(F_\tau^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ is nonempty (cf. Theorem 5). Furthermore, the addition of a monotone term (i.e., $\tau(x - y)$ does not change the convexity of utilities to their actions. We set the vector y to be $y = x(\epsilon^{(j-1)})$, which means that while computing the j-th member of solutions of $VI(F_{\tau,\epsilon}^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$, namely as $x(\epsilon^{(j)})$, the vector y is the same as the solution found for $VI(F_{\tau,\epsilon}^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ when $\epsilon = \epsilon^{(j-1)}$. Therefore, in the limit point where $x_\infty \in \text{SOL}(F_\tau^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$, we have

$$\begin{aligned} &x_\infty \in \text{SOL}(F_\tau^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}) \Rightarrow (x - x_\infty) F_\tau^{\mathbb{R}}(x_\infty) > 0 \\ &\Rightarrow (x - x_\infty) \left(F^{\mathbb{R}}(x_\infty) + \tau(x_\infty - x_\infty) \right) > 0 \\ &\Rightarrow x_\infty \in \text{SOL}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}). \end{aligned} \tag{56}$$

Hence, the term $\tau(x_\infty - x_\infty)$ vanishes since the limit point is guaranteed to be reached.

F. Distributed Tikhonov Regularization

Tikhonov regularization (QNE selection) is done in two nested loops. In the inner loop, for a given $\epsilon^{(j)}$, the solution to $\text{VI}(F_\epsilon^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}})$ will be found from the iteration in (47) (where the multiplier of ϵ is replaced

[16]Later as we proceed, we present the equivalent regularization for the complex version of $F^{\mathbb{R}}$, i.e., $F^{\mathbb{C}}$ as well.

with $\nabla\Phi(x)$). In the outer loop, the next value of $\epsilon^{(j)}$ will be chosen (according to a predefined sequence such that $\lim_{j\to\infty}\epsilon^{(j)}=0$) until the solution to VI$(\nabla\Phi(x),\text{SOL}(F^{\mathbb{R}},\mathcal{K}^{\mathbb{R}}))$ is reached (cf. Theorem 4).

Despite having the ability to select a specific QNE among multiple QNEs, QNE selection requires heavy signaling and centralized computation because still the Lipschitz Continuity constant L and strong monotonicity modulus of $F^{\mathbb{R}}_{\epsilon}(x)$ must be known (cf. Section V-B). In order to address these issues, we introduce another regularization method, namely, proximal point regularization. In this regularization, a term $\theta^{(i)}(x^{(i)}-x^{(i-1)})$ is added to the function $F^{\mathbb{R}}_{\epsilon}(x)$ to build a function $F^{\mathbb{R}}_{\tau,\epsilon,\theta}(x)\triangleq F^{\mathbb{R}}_{\tau,\epsilon}(x)+\theta^{(i)}(x^{(i)}-x^{(i-1)})$ where $\theta^{(i)}$ is a diagonal matrix. Considering this modification, the following property can be used:

Proposition 2. *Let $F^{\mathbb{R}}_{\tau,\epsilon}(x)$ be a strictly monotone and Lipschitz continuous mapping[17]; $\max_{z\in\mathcal{K}^{\mathbb{R}}}||x||\leq C$, and $\max_{z\in\mathcal{K}^{\mathbb{R}}}||F^{\mathbb{R}}_{\tau,\epsilon}||\leq B$ where C and B are positive constants. Furthermore, suppose that for a given $\epsilon^{(j)}$, the solution to VI$(F^{\mathbb{R}}_{\tau,\epsilon},\mathcal{K}^{\mathbb{R}})$ is denoted as $x(\epsilon^{(j)})$. Let $x^{(i)}$ denote the set of iterates defined by*

$$x^{(i+1)}=\Pi_{\mathcal{K}^{\mathbb{R}}}\left(x^{(i)}-\gamma^{(i)}\left(F^{\mathbb{R}}(x^{(i)})+\tau(x^{(i)}-x(\epsilon^{(j-1)}))+\epsilon^{(j)}\nabla\Phi(x^{(i)})+\theta^{(i)}(x^{(i)}-x^{(i-1)})\right)\right) \tag{57}$$

where the step size matrix $\gamma^{(i)}$ is changing with the iterations. Lastly, set $\gamma^{(i)}\theta^{(i)}=c=diag([c_1,\ldots,c_m])$ where $c_{i'}\in(0,1),\forall i'=1,\ldots,m$ is a constant, and let the following hold:

$$\sum_{i=1}\gamma^{(i)}=\infty,\quad \sum_{i=1}^{\infty}\left(\gamma^{(i)}\right)^2<\infty,\quad \textit{and}\ \sum_{i=1}^{\infty}(\gamma^{(i)}_{max}-\gamma^{(i)}_{min})<\infty. \tag{58}$$

where $\gamma^{(i)}_{max}$ and $\gamma^{(i)}_{min}$ are respectively the maximum and minimum diagonal elements of the matrix $\gamma^{(i)}$. Therefore, we have $\lim_{i\to\infty}x^{(i)}=x(\epsilon^{(j)})$. □

The proof of Proposition 2 can be found in [23, Proposition 3.4]. However, note that the assumption of strict monotonicity of $F^{\mathbb{R}}_{\tau,\epsilon}(x)$ is immediately satisfied as $F^{\mathbb{R}}_{\tau,\epsilon}(x)$ is already strongly monotone (cf. (49)). The conditions $\max_{z\in\mathcal{K}^{\mathbb{R}}}||x||\leq C$ and $\max_{z\in\mathcal{K}^{\mathbb{R}}}||F^{\mathbb{R}}_{\tau,\epsilon}||\leq B$ can also be satisfied due to having power constraints on each link. According to [23, Proposition 3.4], the step size $\gamma^{(i)}$ can be chosen as $\gamma^{(i)}_{i'}=(i+\alpha_{i'})^{-\omega}$ where $\alpha_{i'}$ is a positive integer for $i'=1,\ldots,N$ and $0<\omega<1$. Hence, we can write

$$\gamma^{(i)}_{max}=(i+\alpha_{max})^{-\omega},\quad \gamma^{(i)}_{min}=(i+\alpha_{min})^{-\omega}. \tag{59}$$

Note that in Proposition 2, $\theta^{(i)}$ is already set to $\theta^{(i)}=\frac{c}{\gamma^{(i)}}$. Using Proposition 2, we can design a distributed transmit optimization algorithm without the knowledge of Lipschitz constant and strong monotonicity modulus of $F^{\mathbb{R}}_{\tau,\epsilon}$. The next section discusses the implementation of QNE selection using (57)[18].

VI. The QNE Selection Algorithm: Design and Discussion

In this section of the paper, we propose the QNE selection algorithm together with three possible choices for the design criterion (i.e., $\Phi(x)$). Each of these choices imposes a certain amount of signaling overhead as well as a certain amount of improvement on the performance of Algorithm 1 and Algorithm 2.

A. QNE Selection Algorithm

The pseudo-code for the QNE selection algorithm is shown in Algorithm 3. As mentioned previously, it can be seen in Algorithm 3 that the modified game (i.e., QNE selection algorithm) is comprised of two nested loops: outer loop (i.e., line 1), and inner loop (i.e., line 3). In the outer loop the jth member of $\epsilon^{(j)}$'s is selected. In the inner loop, the game s played among the players, and the players update their

[17]Note that Lipschitz continuity of $F^{\mathbb{R}}_{\tau,\epsilon}(x)$ requires both $F^{\mathbb{R}}(x)$ and $\nabla\Phi(x)$ to be Lipschitz continuous. Hence, the proposed choices for $\Phi(x)$ in the next section are all Lipschitz continuous.

[18]Note that in all of the proposed algorithms throughout this paper, it was assumed that at each round of the game, all of the players are maximizing the utilities. This update fashion is also known as Jacobi implementation. The feasibility of implementing the algorithms using other update fashions (e.g., Gauss-Seidel or Asynchronous) can be a subject of future research.

strategies according to (57). The sequence $\epsilon^{(j)}$ must be a decreasing sequence such that $\lim_{j\to\infty}\epsilon^{(j)}=0$. The operation in line 10 of Algorithm 3 can be written as

$$\begin{pmatrix} \mathbf{\Sigma}_q^{(i+1)} \\ \mathbf{W}_q^{(i+1)} \end{pmatrix} = \mathtt{Proj}_{\mathcal{F}_q} \begin{pmatrix} \mathbf{\Sigma}_q^{(i)} + \gamma_q' \left(\nabla_{\mathbf{\Sigma}_q} \bar{f}_q + \tau_q \left(\mathbf{\Sigma}_q^{(i)} - \mathbf{\Sigma}_q(\epsilon^{(j-1)}) \right) + \epsilon^{(j)} \nabla_{\mathbf{\Sigma}_q} \Phi(x^{(i)}) - \theta_q^{(i)} (\mathbf{\Sigma}_q^{(i)} - \mathbf{\Sigma}_q^{(i-1)}) \right) \\ \mathbf{W}_q^{(i)} + \gamma_q' \left(\nabla_{\mathbf{W}_q} \bar{f}_q + \tau_q \left(\mathbf{W}_q^{(i)} - \mathbf{W}_q(\epsilon^{(j-1)}) \right) + \epsilon^{(j)} \nabla_{\mathbf{W}_q} \Phi(x^{(i)}) - \theta_q^{(i)} (\mathbf{W}_q^{(i)} - \mathbf{W}_q^{(i-1)}) \right) \end{pmatrix}. \tag{60}$$

Notice that $\theta_q^{(i)}$ is a diagonal matrix that can obtained via dividing the matrix $\theta^{(i)}$ into Q block-diagonal matrices. That is, (with a slight abuse of notations) $\theta^{(i)} = \text{diag}(\theta_1^{(i)}, \ldots, \theta_Q^{(i)})$. In the next subsection, we specifically explain the terms $\nabla_{\mathbf{\Sigma}_q}\Phi(x)$ and $\nabla_{\mathbf{W}_q}\Phi(x)$ in line 10, so that Algorithm 3 will be completely defined. Lastly, notice that all of our analysis on VI problems were under the assumption that every player is solving a minimization problem as his strategy. Hence, if maximization is the strategy of each player, the proximal terms in (60) appear as a negative values. Furthermore, the addition of $\nabla_{\mathbf{\Sigma}_q}\Phi(x)$ and $\nabla_{\mathbf{W}_q}\Phi(x)$ means that $\Phi(x)$ must be a strongly concave function of x.

Algorithm 3 QNE Selection Algorithm

Initialize: $\mathbf{\Sigma}_q^{(1)}$, $\mathbf{W}_q^{(1)}$, tr($\mathbf{\Sigma}_q^{(1)} + \mathbf{W}_q^{(1)}$) $< P_q$, $\forall q$, and j = 1

1: **repeat** % Outer loop: superscript (j) indicates the iterations starting from here
2: Choose the jth member of the sequence $\epsilon^{(j)}$
3: **repeat** % Inner loop: superscript (i) indicates the iterations starting from here
4: Compute $\mathbf{M}_q$, $\mathbf{M}_{e,q,k}$, $\forall (q,k) \in \mathbb{Q}\times\mathbb{K}$
5: Compute $\mathbf{S}_{q,k}^{(i)}$, $\forall (q,k) \in \mathbb{Q}\times\mathbb{K}$
6: Compute $\varphi_{e,q,k}(\mathbf{\Sigma}_q^{(i)}, \mathbf{W}_q^{(i)}, \mathbf{S}_{q,k}^{(i)})$, $\forall (q,k) \in \mathbb{Q}\times\mathbb{K}$
7: **for** q = 1,...,Q **do**
8: Update the values of τ_q for all $q = 1, \ldots, Q$ such that the inequality in (41) is satisfied
9: Replace $\nabla_{\mathbf{\Sigma}_q}\bar{f}_q$ with $\nabla_{\mathbf{\Sigma}_q}\bar{f}_q - \tau_q \left(\mathbf{\Sigma}_q^{(i)} - \mathbf{\Sigma}_q(\epsilon^{(j-1)}) \right) + \epsilon^{(j)} \nabla_{\mathbf{\Sigma}_q}\Phi(x^{(i)}) - \theta_q^{(i)} \left(\mathbf{\Sigma}_q^{(i)} - \mathbf{\Sigma}_q^{(i-1)} \right)$
10: Replace $\nabla_{\mathbf{W}_q}\bar{f}_q$ with $\nabla_{\mathbf{W}_q}\bar{f}_q - \tau_q \left(\mathbf{W}_q^{(i)} - \mathbf{W}_q(\epsilon^{(j-1)}) \right) + \epsilon^{(j)} \nabla_{\mathbf{W}_q}\Phi(x^{(i)}) - \theta_q^{(i)} \left(\mathbf{W}_q^{(i)} - \mathbf{W}_q^{(i-1)} \right)$
11: Compute $(\mathbf{\Sigma}_q^{(i+1)}, \mathbf{W}_q^{(i+1)})$ using (60)
12: **end for**
13: **until** Convergence to QNE % $x(\epsilon^j)$ is found
14: j = j+1
15: **until** Convergence to limit point of $x(\epsilon^j)$'s

B. On the Choice of Design Criterion for QNE Selection

Assume that the derivatives of $\Phi(x)$ are described as:

$$\nabla\Phi(x) \triangleq [\nabla^{\mathbb{R}}_{\mathbf{\Sigma}_1,\mathbf{W}_1}\Phi(x)^T, \ldots, \nabla^{\mathbb{R}}_{\mathbf{\Sigma}_Q,\mathbf{W}_Q}\Phi(x)^T]^T, \tag{61a}$$

$$\nabla^{\mathbb{R}}_{\mathbf{\Sigma}_q,\mathbf{W}_q}\Phi(x) \triangleq [\nabla^{\mathbb{R}}_{\mathbf{\Sigma}_q}\Phi(x)^T, \nabla^{\mathbb{R}}_{\mathbf{W}_q}\Phi(x)^T]^T, \ q \in \mathbb{Q}, \tag{61b}$$

$$\nabla^{\mathbb{R}}_{\mathbf{\Sigma}_q}\Phi(x) \triangleq [\text{Re}\{\text{vec}(\nabla_{\mathbf{\Sigma}_q}\Phi(x))\}^T, \text{Im}\{\text{vec}(\nabla_{\mathbf{\Sigma}_q}\Phi(x))\}^T]^T, \tag{61c}$$

$$\nabla^{\mathbb{R}}_{\mathbf{W}_q}\Phi(x) \triangleq [\text{Re}\{\text{vec}(\nabla_{\mathbf{W}_q}\Phi(x))\}^T, \text{Im}\{\text{vec}(\nabla_{\mathbf{W}_q}\Phi(x))\}^T]^T. \tag{61d}$$

We are now ready to present the possible choices of $\Phi(x)$:

1) Maximizing the sum of information rates: We aim to select the QNE that maximizes the sum-rate of all links. Recalling the reformulated information rate (i.e., $\varphi_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,k})$) in (10b), $\Phi(x)$ can be

described as (with $q \in \mathbb{Q}$):

$$\nabla_{\mathbf{\Sigma}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{H}_{qr}^H \left((\mathbf{M}_r + \mathbf{H}_{rr}\mathbf{\Sigma}_r\mathbf{H}_{rr}^H)^{-1} - \mathbf{S}_{r,0}\right)\mathbf{H}_{qr}, \tag{62a}$$

$$\nabla_{\mathbf{W}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{H}_{qr}^H \left((\mathbf{M}_r + \mathbf{H}_{rr}\mathbf{\Sigma}_r\mathbf{H}_{rr}^H)^{-1} - \mathbf{S}_{r,0}\right)\mathbf{H}_{qr}. \tag{62b}$$

Notice that although we wrote Φ as a function of x, one can easily relate the vector x to the covariance matrices $\{(\mathbf{\Sigma}_q, \mathbf{W}_q)\}_{q=1}^{Q}$ using (61) and (43). Hence, the derivatives of $\Phi(x)$ at the end of Algorithm 3 would be:

$$\nabla_{\mathbf{\Sigma}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{H}_{qr}^H \left((\mathbf{M}_r^\star + \mathbf{H}_{rr}\mathbf{\Sigma}_r^\star\mathbf{H}_{rr})^{-1} - \mathbf{S}_{r,0}^\star\right)\mathbf{H}_{qr} \tag{63a}$$

$$\nabla_{\mathbf{W}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q} \mathbf{H}_{qr}^H \left((\mathbf{M}_r^\star + \mathbf{H}_{rr}\mathbf{\Sigma}_r^\star\mathbf{H}_{rr})^{-1} - \mathbf{S}_{r,0}^\star\right)\mathbf{H}_{qr} \tag{63b}$$

where $\mathbf{M}_r^\star = \mathbf{I} + \mathbf{H}_{rr}(\mathbf{W}_r^\star)\mathbf{H}_{rr}^H + \mathbf{H}_{qr}(\mathbf{W}_q^\star + \mathbf{\Sigma}_q^\star)\mathbf{H}_{qr}^H + \sum_{\substack{l=1\\l\neq q,r}}^{Q}\mathbf{H}_{lr}\left(\mathbf{\Sigma}_l^\star + \mathbf{W}_l^\star\right)\mathbf{H}_{lr}^H$, with $\mathbf{\Sigma}_q^\star$ and $\mathbf{W}_q^\star$ being the limit points of $\mathbf{\Sigma}_q$ and $\mathbf{W}_q$. Integrating (63a) w.r.t. $\mathbf{\Sigma}_q^\star$ and integrating (63b) w.r.t. $\mathbf{W}_q^\star$, we end up with $\Phi(x) = \sum_{q=1}^{Q}\sum_{\substack{r=1\\r\neq q}}^{Q}\varphi_r(\mathbf{\Sigma}_r, \mathbf{W}_r, \mathbf{S}_{r,0})$. Hence, at the end of Algorithm 3, the QNE that is a stationary point of sum-rate of all links is selected, i.e., the point that is the unique solution of $\mathrm{VI}(\nabla\Phi(x), \mathrm{SOL}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}))$.

2) Minimizing the received rates at Eves: We can describe $\Phi(x)$ by (with $q \in \mathbb{Q}$)

$$\nabla_{\mathbf{\Sigma}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q}\sum_{k=1}^{K}\rho_{r,k}\mathbf{G}_{rk}^H\left({\mathbf{M}_{e,r,k}}^{-1} - \mathbf{S}_{r,k}\right)\mathbf{G}_{rk} \tag{64a}$$

$$\nabla_{\mathbf{W}_q}\Phi(x) = \sum_{\substack{r=1\\r\neq q}}^{Q}\sum_{k=1}^{K}\rho_{r,k}\mathbf{G}_{rk}^H\left({\mathbf{M}_{e,r,k}}^{-1} - \mathbf{S}_{r,k}\right)\mathbf{G}_{rk} \tag{64b}$$

$$\mathbf{M}_{e,r,k} \triangleq \mathbf{I} + \mathbf{G}_{rk}\mathbf{W}_r\mathbf{G}_{rk}^H + \mathbf{G}_{qk}\left(\mathbf{\Sigma}_q + \mathbf{W}_q\right)\mathbf{G}_{qk}^H + \tag{64c}$$

$$\sum_{\substack{l=1\\l\neq q,r}}^{Q}\mathbf{G}_{lk}\left(\mathbf{\Sigma}_l + \mathbf{W}_l\right)\mathbf{G}_{lk}^H \tag{64d}$$

where the term $\rho_{r,k}$ is defined in (72). Following the same reasoning used in the previous QNE selection, at the limit point of $x(\epsilon^{(j)})$, we end up with $\Phi(x) = \sum_{q=1}^{Q}\sum_{\substack{r=1\\r\neq q}}^{Q} -\frac{1}{\beta}\ln(\sum_{k=1}^{K}\exp\{\beta\varphi_{e,r,k}(\mathbf{\Sigma}_r, \mathbf{W}_r, \mathbf{S}_{r,k})\})$, where $\varphi_{e,r,k}(\mathbf{\Sigma}_r, \mathbf{W}_r, \mathbf{S}_{r,k})$ is defined in (10c). Hence, the selected QNE guides the game to the stationary point of minimizing Eves' received rates, i.e., the point that is the unique solution of $\mathrm{VI}(\nabla\Phi(x), \mathrm{SOL}(F^{\mathbb{R}}, \mathcal{K}^{\mathbb{R}}))$.

3) Maximizing the sum of secrecy rates: In this criterion, a simple addition of previous design criteria gives us another QNE selection method, in which the QNE that is a stationary point of secrecy sum-rate is selected.

VII. Centralized Algorithm

An appropriate measure of efficiency (i.e., social welfare) in our game would be the sum of utilities of all players or the secrecy sum-rate. The price of anarchy (PoA) can be defined as the ratio between

the performance of the optimal centralized solution for the secrecy sum-rate maximization problem and the *worst* NE. However, such definition of PoA requires us to solve the secrecy sum-rate maximization problem, which is a nonconvex problem. Moreover, all of the proposed algorithms converge to the QNEs of the proposed game, which are not necessarily NEs. Hence, direct PoA analysis is not feasible. Instead, to measure the efficacy of QNEs, we design a centralized algorithm that provides locally optimal solutions for the (network-wide) secrecy sum-rate maximization problem. We refer to this algorithm as Centralized Secrecy Sum-rate Maximization method (CSSM). The objective value of solving (66) via the CSSM is considered as the social welfare in our game.

In CSSM, the objective is to find a stationary solution for the following optimization problem:

$$\underset{(\mathbf{\Sigma}_q,\mathbf{W}_q)\in\mathcal{F}_q,\ \forall q}{\text{maximize}} \sum_{q=1}^{Q} \bar{R}_{s,q}(\mathbf{\Sigma}_q,\mathbf{W}_q). \tag{65}$$

Using the reformulation techniques given in Section III, the problem in (65) can be rewritten as

$$\begin{aligned}\underset{\substack{(\mathbf{\Sigma}_q,\mathbf{W}_q,\mathbf{S}_{q,k})\\ \forall q,k}}{\text{maximize}} \ & \sum_{q=1}^{Q} \bar{f}_q(\mathbf{\Sigma}_q,\mathbf{W}_q,\{\mathbf{S}_{q,k}\}_{k=0}^{K}) \\ \text{s.t.} \ & (\mathbf{\Sigma}_q,\mathbf{W}_q)\in\mathcal{F}_q,\forall q\in\mathbb{Q}, \\ & \mathbf{S}_{q,k}\succeq 0,\ \forall (q,k)\in\mathbb{Q}\times\{0\}\cup\mathbb{K}. \end{aligned} \tag{66}$$

Problem (66) can be shown to be convex w.r.t either $[\mathbf{\Sigma},\mathbf{W}]=\{\mathbf{\Sigma}_q,\mathbf{W}_q\}_{q=1}^{Q}=\left[[\mathbf{\Sigma}_1,\mathbf{W}_1]^T,\ldots,[\mathbf{\Sigma}_Q,\mathbf{W}_Q]^T\right]$ or $\mathbf{S}=\{\mathbf{S}_{q,k}\}_{\forall q,k}=[\mathbf{S}_{1,0},\ldots,\mathbf{S}_{1,K},\mathbf{S}_{2,0},\ldots,\mathbf{S}_{2,K},\ldots,\mathbf{S}_{Q,K}]^T$. Hence, a stationary point of (65) can be found by solving (66) sequentially w.r.t. $[\mathbf{\Sigma},\mathbf{W}]$ and $\mathbf{S}$ until reaching a convergence point. That is, in one iteration, problem (66) is solved w.r.t. $\mathbf{S}$ to find an optimal solution $\mathbf{S}^*$. Next, with $\mathbf{S}^*$ plugged in the objective of (66), problem (66) can be optimized w.r.t. $[\mathbf{\Sigma},\mathbf{W}]$ to find an optimal solution $[\mathbf{\Sigma}^*,\mathbf{W}^*]=\{\mathbf{\Sigma}_q^*,\mathbf{W}_q^*\}_{q=1}^{Q}$. Problem (66) is separable w.r.t. every element of $\mathbf{S}$. Hence,

$$\mathbf{S}_{q,0}^* \triangleq \arg\max_{\mathbf{S}_{q,0}\succeq 0} \sum_{q=1}^{Q} \bar{f}_q(\mathbf{\Sigma}_q,\mathbf{W}_q,\mathbf{S}_{q,k}) = (\mathbf{M}_q)^{-1}$$

$$\mathbf{S}_{q,k}^* \triangleq \arg\max_{\mathbf{S}_{q,k}\succeq 0} \sum_{q=1}^{Q} \bar{f}_q(\mathbf{\Sigma}_q,\mathbf{W}_q,\mathbf{S}_{q,k}) =$$

$$\left(\mathbf{M}_{e,q,k}+\mathbf{G}_{qk}\mathbf{\Sigma_q}^{n}\mathbf{G}_{qk}^H\right)^{-1} =$$

Now, we can solve (66) w.r.t $[\mathbf{\Sigma},\mathbf{W}]$. We use the augmented Lagrangian multiplier method [24] to derive a centralized solution for $[\mathbf{\Sigma}^*,\mathbf{W}^*]$. Let $\mathbf{c}_q=\text{tr}(\mathbf{\Sigma}_q+\mathbf{W}_q)-P_q<0$. The augmented Lagrangian of (66) is [24][19]

$$\begin{aligned} L(\mathbf{\Sigma},\mathbf{W},\mathbf{a},\mathbf{p},\mathbf{S}^*) = & -\sum_{q=1}^{Q} \bar{f}_q(\mathbf{\Sigma}_q,\mathbf{W}_q,\{\mathbf{S}_{q,k}^*\}_{k=0}^{K}) + \\ & \frac{1}{2\mathbf{p}}\sum_{q=1}^{Q}\left\{(\max\{\mathbf{a}_q+\mathbf{p}\mathbf{c}_q,0\})^2+\mathbf{a}_q^2\right\} \end{aligned} \tag{68}$$

where $\mathbf{p}$ is a positive penalty (to prevent constraint violations) and $\mathbf{a}_q$, $q=1,\ldots,Q$, are the nonnegative

[19]We converted the problem in (66) to a minimization problem by considering the negative of the objective function.

Lagrange multipliers. At a stationary point, the following equalities hold for all $q \in \mathbb{Q}$

$$\frac{\partial}{\partial \mathbf{\Sigma}_q} L(\mathbf{\Sigma}, \mathbf{W}, \mathbf{a}, \mathbf{p}, \mathbf{S}^*) = -\sum_{r=1}^{Q} \frac{\partial}{\partial \mathbf{\Sigma}_q} \bar{f}_r(\mathbf{\Sigma}_r, \mathbf{W}_r, \{\mathbf{S}_{r,k}^*\}_{k=0}^K) + \frac{1}{2\mathbf{p}} \sum_{r=1}^{Q} \frac{\partial}{\partial \mathbf{\Sigma}_q} (\max\{\mathbf{a}_r + \mathbf{pc}_r, 0\})^2 = 0 \tag{69a}$$

$$\frac{\partial}{\partial \mathbf{W}_q} L(\mathbf{\Sigma}, \mathbf{W}, \mathbf{a}, \mathbf{p}, \mathbf{S}^*) = -\sum_{r=1}^{Q} \frac{\partial}{\partial \mathbf{W}_q} \bar{f}_r(\mathbf{\Sigma}_r, \mathbf{W}_r, \{\mathbf{S}_{r,k}^*\}_{k=0}^K) + \frac{1}{2\mathbf{p}} \sum_{r=1}^{Q} \frac{\partial}{\partial \mathbf{W}_q} (\max\{\mathbf{a}_r + \mathbf{pc}_r, 0\})^2 = 0 \tag{69b}$$

where

$$\frac{\partial}{\partial \mathbf{\Sigma}_q} \bar{f}_r(\mathbf{\Sigma}_r, \mathbf{W}_r, \{\mathbf{S}_{r,k}^*\}_{k=0}^K) = \begin{cases} \mathbf{H}_{qq}^H (\mathbf{M}_q + \mathbf{H}_{qq} \mathbf{\Sigma}_q \mathbf{H}_{qq}^H)^{-1} \mathbf{H}_{qq} - \sum_{k=1}^{K} \rho_{q,k} \mathbf{G}_{q,k}^H \mathbf{S}_{q,k}^* \mathbf{G}_{q,k}, & r = q, \\ \mathbf{H}_{qr}^H \left((\mathbf{M}_r + \mathbf{H}_{rr} \mathbf{\Sigma}_r \mathbf{H}_{rr}^H)^{-1} - \mathbf{S}_{r,0}^* \right) \mathbf{H}_{qr} + \sum_{k=1}^{K} \rho_{r,k} \mathbf{G}_{rk}^H \left((\mathbf{M}_{e,r,k}{}^{-1} - \mathbf{S}_{r,k}^* \right) \mathbf{G}_{rk}, & r \neq q \end{cases} \tag{70}$$

$$\frac{\partial}{\partial \mathbf{W}_q} \bar{f}_r(\mathbf{\Sigma}_r, \mathbf{W}_r, \{\mathbf{S}_{r,k}^*\}_{k=0}^K) = \begin{cases} \mathbf{H}_{qq}^H \left((\mathbf{M}_q + \mathbf{H}_{qq} \mathbf{\Sigma}_q \mathbf{H}_{qq})^{-1} - \mathbf{S}_{q,0}^* \right) \mathbf{H}_{qq} + \sum_{k=1}^{K} \rho_{q,k} \mathbf{G}_{qk}^H \left((\mathbf{M}_{e,q,k})^{-1} - \mathbf{S}_{q,k}^* \right) \mathbf{G}_{qk}, & r = q, \\ \mathbf{H}_{qr}^H \left((\mathbf{M}_r + \mathbf{H}_{rr} \mathbf{\Sigma}_r \mathbf{H}_{rr}^H)^{-1} - \mathbf{S}_{r,0}^* \right) \mathbf{H}_{qr} + \sum_{k=1}^{K} \rho_{r,k} \mathbf{G}_{rk}^H \left((\mathbf{M}_{e,r,k})^{-1} - \mathbf{S}_{r,k}^* \right) \mathbf{G}_{rk}, & r \neq q, \end{cases} \tag{71}$$

and

$$\rho_{q,k} = \frac{e^{\beta \varphi_{e,q,k}(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,k}^*)}}{\sum_{j=1}^{K} e^{\beta \varphi_{e,q,j}(\mathbf{\Sigma}_q, \mathbf{W}_q, \mathbf{S}_{q,j}^*)}}. \tag{72}$$

The second term in the RHS of (69a) is continuously differentiable w.r.t $\mathbf{\Sigma}_q$ when $r = q$ [24, pp. 397]. Thus,

$$\frac{\partial}{\partial \mathbf{\Sigma}_q} (\max\{\mathbf{a}_r + \mathbf{pc}_r, 0\})^2 = \begin{cases} 2\mathbf{p}(\mathbf{a}_q + \mathbf{pc}_q) \mathbf{\Sigma}_q, r = q, \mathbf{a}_q + \mathbf{pc}_q > 0 \\ 0, \qquad \text{ow.} \end{cases} \tag{73}$$

and

$$\frac{\partial}{\partial \mathbf{W}_q} (\max\{\mathbf{a}_r + \mathbf{pc}_r, 0\})^2 = \begin{cases} 2\mathbf{p}(\mathbf{a}_q + \mathbf{pc}_q) \mathbf{W}_q, r = q, \mathbf{a}_q + \mathbf{pc}_q > 0 \\ 0, \qquad \text{ow.} \end{cases} \tag{74}$$

To satisfy the conditions in (69), we used gradient descent with a line search satisfying Armijo rule. The details of the centralized algorithm is presented in Algorithm CSSM. The centralized nature of Algorithm CSSM can be seen in Line 12, where the equalities in (69) are checked for all $q \in \mathbb{Q}$ and Line 11, where the Armijo rule is applied. The convergence of this algorithm can be proved by extending the proof of Theorem 1 in the paper and [25, Corollary 2], which is skipped here for the sake of brevity. Note that Algorithm CSSM is sensitive to the initial values of $[\mathbf{\Sigma}, \mathbf{W}]$. Thus, we simulated this algorithm with random initializations and averaged its performance over the total number of initializations.

Algorithm CSSM The Centralized Secrecy Sum-rate Maximization Algorithm (CSSM)

Initialize: $\mathbf{\Sigma}_q^{(1)}$, $\mathbf{W}_q^{(1)}$, $\text{tr}(\mathbf{\Sigma}_q^{(1)} + \mathbf{W}_q^{(1)}) < P_q$, $\forall q$, $i = 0$

repeat i = i+1 % superscript (i) indicates the iterations starting from here.
  Compute $\mathbf{S}_{q,k}^{(i)}$, $\forall (q,k) \in \mathbb{Q} \times \mathbb{K}$, $\mathbf{p} = 1$, $\mathbf{a}_q = 0,\ \forall q$, and s_t (Armijo step size)
  repeat Set $m = 1$
    repeat Set $n = 1$ % superscript (m) indicates the iterations starting from here.
      Set$[d_{\mathbf{\Sigma}_q}, d_{\mathbf{W}_q}]^T = -[\frac{\partial}{\partial \mathbf{\Sigma}_q} L^{(m)^T}, \frac{\partial}{\partial \mathbf{W}_q} L^{(m)^T}],\ \forall q \Rightarrow d = \{d_{\mathbf{\Sigma}_q}, d_{\mathbf{W}_q}\}_{q=1}^{Q}$
      Set $[\hat{\mathbf{\Sigma}}, \hat{\mathbf{W}}] = [\mathbf{\Sigma^{(m)}}, \mathbf{W^{(m)}}] + \mathbf{d}$
      Set $[\mathbf{\Sigma^{(m+1)}}, \mathbf{W^{(m+1)}}] = [\mathbf{\Sigma^{(m)}}, \mathbf{W^{(m)}}] + \mathbf{s_t^n}([\hat{\mathbf{\Sigma}}, \hat{\mathbf{W}}] - [\mathbf{\Sigma^{(m)}}, \mathbf{W^{(m)}}])$
      repeat % superscript (n) indicates the iterations starting from here.
        $s_t^{n+1} = s_t(s_t^n)$
        Set $[\mathbf{\Sigma^{(m+1)}}, \mathbf{W^{(m+1)}}] = [\mathbf{\Sigma^{(m)}}, \mathbf{W^{(m)}}] + \mathbf{s_t^{n+1}}([\hat{\mathbf{\Sigma}}, \hat{\mathbf{W}}] - [\mathbf{\Sigma^{(m)}}, \mathbf{W^{(m)}}])$
      until $L(\mathbf{\Sigma}^{(m+1)}, \mathbf{W}^{(m+1)}, \mathbf{a}^{(m+1)}, \mathbf{p}, \mathbf{S}^{(i)}) < L(\mathbf{\Sigma}^{(m)}, \mathbf{W}^{(m)}, \mathbf{a}^{(m)}, \mathbf{p}, \mathbf{S}^{(i)}) + s_t^n d^T \{\frac{\partial}{\partial \mathbf{\Sigma}_q} L^{(m)}, \frac{\partial}{\partial \mathbf{W}_q} L^{(m)}\}_q$
    until $\frac{\partial}{\partial \mathbf{\Sigma}_q} L = \frac{\partial}{\partial \mathbf{W}_q} L = 0,\ \forall q$
    $\mathbf{a}_q = \max\{\mathbf{a}_q + p\mathbf{c}_q, 0\}$
    $\mathbf{p} = \mathbf{p} \times u$ % $u \geq 1$ increase the penalty.
  until $\max\{\mathbf{c}_1, \ldots, \mathbf{c}_q\} \leq 0$
until Convergence of $L(\mathbf{\Sigma}, \mathbf{W}, \mathbf{a}, \mathbf{p}, \mathbf{S})$

VIII. Simulation Results and Discussion

In this section, we simulate and compare all the algorithms presented so far. In these simulations, we set the noise power to 0 dBm. Q links as well as K eavesdroppers are randomly placed in a circle, namely, the simulation region, with radius r_{circ}. The distance between the transmitter and the receiver of each link is set to be a constant $d_{link} = 10$ m. The path-loss exponent is set to 2.5. For all simulated algorithms, $\beta = 5$ (cf. (12)) and the termination criterion is set to when the normalized relative difference in each link's secrecy rate for two consecutive iterations is less than 10^{-3}. For the QNE selection algorithms, we set their parameters as follows: The step size matrix (i.e., γ') is set such that $\gamma'^{(i)}_j = \gamma_0 i^{(-0.6)}$, $j = 1, \ldots, m$, where γ_0 is a positive constant[20], $c = 0.08 I_{m \times m}$, and $\epsilon^{(j)} = \frac{1}{j}$.

A. Signaling Overhead and Running Time

While the distributed implementation of our proposed algorithms is now complete (cf. (45) and (60)), we still need to make sure that the amount of coordination that each link has to do (to make each QNE selection method possible) is reasonably low. That is, we need to check how much (if any) information a link needs to know about other links' corresponding channels and transmission attributes (i.e., covariance matrices of information signal and AN) in order to execute one iteration of each algorithm.

Algorithm 1 presented in the previous manuscript only requires each link to measure the interference at its receiver to perform the optimization in (51). By examining the iteration in (45) for each link, where $\nabla_{\mathbf{\Sigma}_q} \bar{f}_q$ and $\nabla_{\mathbf{W}_q} \bar{f}_q$ are given in (21), we can deduce that Algorithm 2 requires the same amount of coordination as Algorithm 1. The amount of coordination for Algorithm 3, however, depends on the choice of the function $\Phi(x)$. Here, we compare all of the flavors of Algorithm 3 in terms of how much signaling overhead they impose on the network. If maximizing sum-rate is the criterion, from (62) it can be seen that during the computation of $x(\epsilon^{(j)})$, at each iteration, the qth link, $q \in \mathbb{Q}$, needs the values of received signal, noise-plus-interference, and $\mathbf{S}_{r,0}$ ($r \in \mathbb{Q}, r \neq q$) of other links. Furthermore, the cross-channel gains of the qth link with other (unintended) legitimate receivers (i.e., $\mathbf{H}_{qr}, \forall r \in \mathbb{Q}, r \neq q$) should also be

[20]We found out that setting the maximum value of $\gamma_0 = 20000$ brings the best performance for our algorithms.

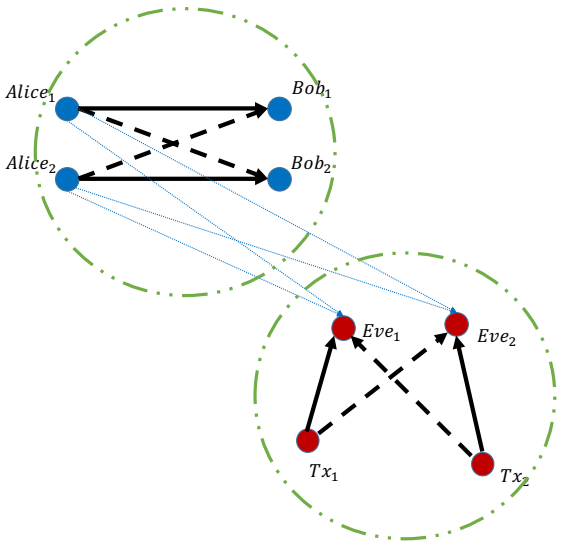


Fig. 2: A (clustered) MANET where two clusters (indiacted by green circle) of ad-hoc nodes are near each other. Hence, one cluster might be interested in the ongoing communications of the other cluster.

available. Note that the cross-channel gains need not to be acquired multiple times at each iteration, as they are fixed throughout the coherence time of the channels[21]. If the rth receiver sends training signals to its corresponding transmitter, for (implicit) channel estimation, $r \in \mathbb{Q}, r \neq q$, the channel gains $\mathbf{H}_{qr}$ can be estimated by the qth transmitter using channel reciprocity. Moreover, it should be noted that while the qth link, $q = 1, \ldots, Q$, is using this criterion, it does not need to know any information about the channel gains between other links and eavesdroppers. This feature makes this design criterion more favorable than other criteria, which require obtaining the eavesdropping channel gains (i.e., $\mathbf{G}_{rk}$ and $\mathbf{S}_{r,k}, \forall r \neq q, \forall k$) of all other links.

For the case of passive eavesdroppers, it does not seem difficult to derive the responses (or gradients) while assuming the knowledge of only statistics of the eavesdropping channels. This can be done if in (16) we replace the term $\varphi_{e,q,k}$ with $E\left[\varphi_{e,q,k}\right]$ where the expectation is w.r.t $G_{qk}, \forall q, k \in \mathbb{Q} \times \mathbb{K}$. Note that including the expectation operator in the utilities, does not compromise the generality of any of the analyses done in previous sections. Despite general difficulties in acquiring eavesdroppers's CSI (ECSI), some applications can be considered as practical examples where the knowledge of ECSI can be easily captured. One such example is mobile ad-hoc networks (MANETs) where the ad-hoc links of one cluster are interfering with one another, and can be considered as the legitimate links of our setup (See Fig. 2). On the other hand, the receivers of another cluster may try to overhear the communications of the legitimate links in the nearby clusters. These receivers can be considered as the external eavesdroppers of our setup. The clustering may have been done to ease the routing process in the network. It is possible that the clustering algorithm requires the links to exchange their location, power, and (possibly) channel state information (CSI). Hence, provided that the coherence time of the channels are long enough, each link can maintain the CSI between itself and the links from another cluster. Hence, the ECSI can be known to the links.

Another instance of our setup involves the downlink scenario of current cellular networks. Specifically, assume that the communication of the BS of a cell is interfering with other nearby cells. Each BS-user pair can be assumed as a legitimate link in our scenario. We assume that no MU-MIMO technique is done in this scenario, so a BS is only communicating with one receiver (i.e., UE) at a given time. There might be other idle users in such network that are interested in overhearing the current communications. We can consider these idle users as the external eavesdroppers. It is possible that during the cell association phase, the idle users –which are now the external eavesdroppers– exchange their location information (using

[21]Note that all aforementioned algorithms must run during the coherence time of the channels.

known packets) with all the nearby BSs to eventually select a cell for their respective communications. Hence, the BSs can extract the CSI between themselves and the external eavesdroppers and maintain it (till the end of one coherence time) for use in PHY-layer security optimizations.

The issue of knowledge of ECSI has also been investigated in the recent literature. One example is when Eve is acting as a reactive jammer. That is, after some eavesdropping on the current transmissions, Eve injects her jamming signal to disrupt the ongoing communications. In such a case when jamming happens, assuming that the jamming signal of Eves are previously known, the ECSI can be extracted by the legitimate links using channel reciprocity. Moreover, in [26], it was shown that in a massive MIMO scenario, a passive Eve cannot be very dangerous and must therefore be active and attack the training phase. This active attack can make Eve exposed, and hence the legitimate links can acquire some knowledge about ECSI. Recently, the authors in [27] proposed a method with which the legitimate nodes can detect the passive eavesdropper from the local oscillator power leaked from its RF front end. Hence, an approximation on the location of Eve can be acquired. Lastly, in some scenarios where the legitimate nodes can detect the transmissions from Eves (e.g., active eavesdropping attacks), blind channel estimation techniques can be exploited to capture ECSI [28], [29].

Lastly, regarding the computation of the proximal term τ_q as described by (54), through numerous simulations we found that regardless of the topology of the network and the channel gains, the value found for τ_q is always a vary small value (i.e., $\tau_q < 10^{-4}$). This does not compromise the validity of inequality (54). However, in practice it seems that the transmit optimization game is always a monotone VI problem. The derivation of inequality (54) was done because of the fact that it is not that obvious to see the monotonicity of $VI(F^{\mathbb{C}}, \mathcal{K})$.

It is also interesting to understand how the choice of design criterion changes the running time of our proposed algorithm. To do this, we start from analyzing the computational complexity of Algorithm 1 and extend it to the analysis of our proposed algorithms.

1) Algorithm 1: In Line 2 of Algorithm 1, there is no need to compute every term of $\mathbf{M}_q$ and $\mathbf{M}_{e,q,k}$; that is, in measuring the interference, only the aggregate value is needed. Hence, the complexity of Line 2 is equivalent to the complexity of calculating the covariance matrix of the received interference. More specifically, at the receivers of legitimate links, the covariance matrix calculation of the $N_{R_q} \times 1$ received interference vector (i.e., $\mathbf{M}_q$) yields a complexity of $\mathcal{O}(N_{R_q}^2)$. Similar computation is needed to obtain $\mathbf{M}_{e,q,k}$, which has the complexity of $\mathcal{O}(\sum_{k=1}^{K} N_{ek}^2)$. Line 5 of Algorithm 1 involves a matrix inversion for $\mathbf{S}_{q,0}$ and a matrix multiplication together with a matrix inversion for $\{\mathbf{S}_{q,k}\}_{k=1}^{K}$. The total complexity of this line is $\mathcal{O}(\sum_{k=1}^{K}(N_{T_q}N_{e,k}^2 + N_{e,k}N_{T_q}^2 + N_{e,k}^3) + N_{R_q}^3)$. Computation of the gradients in Line 8 requires the computation of $\varphi_{e,q,k}$ for all $k \in \mathbb{K}$ and $(\mathbf{M}_q^{n,l} + \mathbf{H}_{qq}\mathbf{\Sigma}_q\mathbf{H}_{qq}^H)^{-1}$. Computation of $\varphi_{e,q,k}$ for all $k \in \mathbb{K}$ has the complexity of $\mathcal{O}(\sum_{k=1}^{K} N_{T_q}N_{e,k}^2 + N_{e,k}N_{T_q}^2 + N_{e,k}^3)$ due to matrix multiplications and determinant calculations (cf. (10c)). The inverse of $(\mathbf{M}_q^{n,l} + \mathbf{H}_{qq}\mathbf{\Sigma}_q\mathbf{H}_{qq}^H)$ yields an additional complexity of $\mathcal{O}(N_{R_q}^3 + N_{R_q}^2 N_{T_q} + N_{R_q}N_{T_q}^2)$. Notice that in calculating $\mathbf{M}_q^{n,l}$ and $\mathbf{M}_{e,q,k}^{n,l}$ for all $k \in \mathbb{K}$, an additional computation for calculating $\mathbf{H}_{qq}\mathbf{W}_q^{n,l}\mathbf{H}_{qq}^H$ and $\mathbf{G}_{qk}\mathbf{W}_q^{n,l}\mathbf{G}_{qk}^H$ must be carried at each iteration of the PG method (i.e., Line 6 of Algorithm 1), which respectively have complexities of $\mathcal{O}(N_{R_q}^2 N_{T_q} + N_{R_q}N_{T_q}^2)$ and $\mathcal{O}(\sum_{k=1}^{K} N_{T_q}N_{e,k}^2 + N_{e,k}N_{T_q}^2))$. The other computations that were not mentioned in gradient derivation are redundant and do not affect the general complexity. Apart from the gradient derivations, the Euclidean projection also has its own complexity. The projection in (20) requires eigenvalue decomposition, and thus has $\mathcal{O}(N_{T_q}^3)$ complexity. Adding all of the aforementioned computations, the complexity of Algorithm 1 for each user q is $\mathcal{O}\left(N_{R_q}^3 + N_{T_q}^3 + N_{R_q}^2 N_{T_q} + N_{R_q}N_{T_q}^2 + K(N_{T_q}N_{e,k}^2 + N_{e,k}N_{T_q}^2 + N_{e,k}^3)\right)$ or simply $\mathcal{O}\left(N_{R_q}^3 + N_{T_q}^3 + KN_{e,k}^3\right)$. Note that one might also multiply this complexity by the amount of iterations in the PG method and the AO process. Let the constants N_{PG} and N_{AO} denote the iterations taken

in the PG method and AO process, respectively. Hence, the total complexity for each player q is[22] $\mathcal{O}\left(N_{PG}N_{AO}\left(N_{R_q}^3+N_{T_q}^3+KN_{e,k}^3\right)\right)$.

Algorithm 2: This algorithm can also be handled with the same complexity as Algorithm 1 with the difference that the number of iterations in Algorithm 2 (i.e., repeating the loop at Line 1 of Algorithm 2) was shown in Fig. 5 (a) to be more than Algorithm 1, and hence a slower algorithm compared to Algorithm 1. Let the convergence time of the loop in Line 1 of Algorithm 2 be N_{GR}. Thus, the total complexity of Algorithm 2 for each player q is $\mathcal{O}\left(N_{GR}\left(N_{R_q}^3+N_{T_q}^3+KN_{e,k}^3\right)\right)$.

Algorithm 3: In this algorithm, some additional calculations are generally required. For the criterion of sum-rate maximization, the derivation of the gradients of $\Phi(x)$ are shown in (62), which has the additional complexity of $\mathcal{O}(\sum_{\substack{r=1\\r\neq q}}^{Q}N_{R_r}^3+N_{R_r}^2N_{T_r}+N_{R_r}N_{T_r}^2)$. In the case of minimizing Eves' rates as the QNE selection method, according to (64), computing $\Phi(x)$ would have the complexity of $\mathcal{O}(\sum_{\substack{r=1\\r\neq q}}^{Q}\sum_{k=1}^{K}N_{T_r}N_{e,k}^2+N_{e,k}N_{T_r}^2+N_{e,k}^3)$. The convergence time of Algorithm 3 is generally different from that of Algorithm 2 due to the presence of criterion function in Algorithm 3. Setting N_{QNE} as the convergence time of the loop in Line 1 of Algorithm 3, the total complexity of Algorithm 3 is obtained as follows:

- Under sum-rate maximization as the QNE selection method, for every player q, the computational complexity is

$$\mathcal{O}\left(N_{QNE}N_{GR}\left(N_{T_q}^3+Q(N_{R_q}^3+N_{R_q}^2N_{T_q}+N_{R_q}N_{T_q}^2)+K(N_{T_q}N_{e,k}^2+N_{e,k}N_{T_q}^2+N_{e,k}^3)\right)\right),$$

  or simply

$$\mathcal{O}\left(N_{QNE}N_{GR}\left(N_{T_q}^3+QN_{R_q}^3+KN_{e,k}^3\right)\right). \tag{75}$$

- Under the minimization of Eves' rates as the QNE selection method, for every player q, the complexity is

$$\mathcal{O}\left(N_{QNE}N_{GR}\left(N_{T_q}^3+N_{R_q}^3+N_{R_q}^2N_{T_q}+N_{R_q}N_{T_q}^2+QK(N_{T_q}N_{e,k}^2+N_{e,k}N_{T_q}^2+N_{e,k}^3)\right)\right),$$

  or simply

$$\mathcal{O}\left(N_{QNE}N_{GR}\left(N_{T_q}^3+N_{R_q}^3+QKN_{e,k}^3\right)\right) \tag{76}$$

- Under the maximization of the secrecy sum-rate as the QNE selection method, for every player q, the complexity is

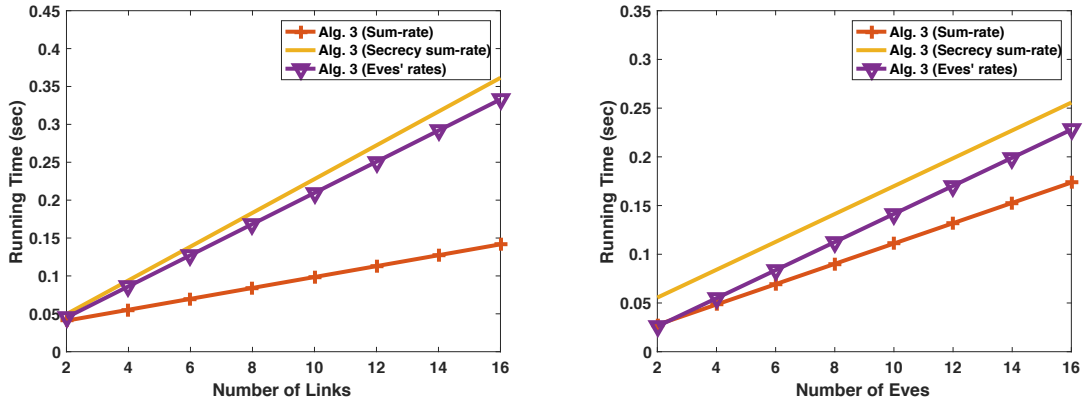
$$\mathcal{O}\left(N_{QNE}N_{GR}Q\left(N_{R_q}^3+N_{T_q}^3+N_{R_q}^2N_{T_q}+N_{R_q}N_{T_q}^2+K(N_{T_q}N_{e,k}^2+N_{e,k}N_{T_q}^2+N_{e,k}^3)\right)\right),$$

  or simply

$$\mathcal{O}\left(N_{QNE}N_{GR}Q\left(N_{R_q}^3+N_{T_q}^3+KN_{e,k}^3\right)\right) \tag{77}$$

We also computed the actual running time of our algorithm using MATLAB on a commercial PC with the following specifications: 1) 2.4 GHz Intel Core i5 CPU, 2) 8 GB 1333 MHz DDR3 RAM, 3) Mac OS X El Capitan v. 10.11.6. We show the results in Fig. 3 for one iteration of Algorithm 3 while using different criteria. Hence, in comparing these results with the theoretical derivations, one should skip the term N_{GR} and N_{QNE}. Each point in the presented curves is averaged over the number of iterations and also over

[22]Notice that this result only makes sense when the QNE is unique. Otherwise if QNE is not unique, Algorithm 2 might not even converge, taking the running time to infinity.



(a) (b)

Fig. 3: (a) Comparison of actual running time of proposed algorithms vs. (a) number of links, and (b) number of Eves: $r_{circ} = 30$ m, $K = 5, N_{T_q} = 5,\ N_{r_q} = 5, N_{e,k} = 5, P_q = 40$ dBm, $\forall q, k$.

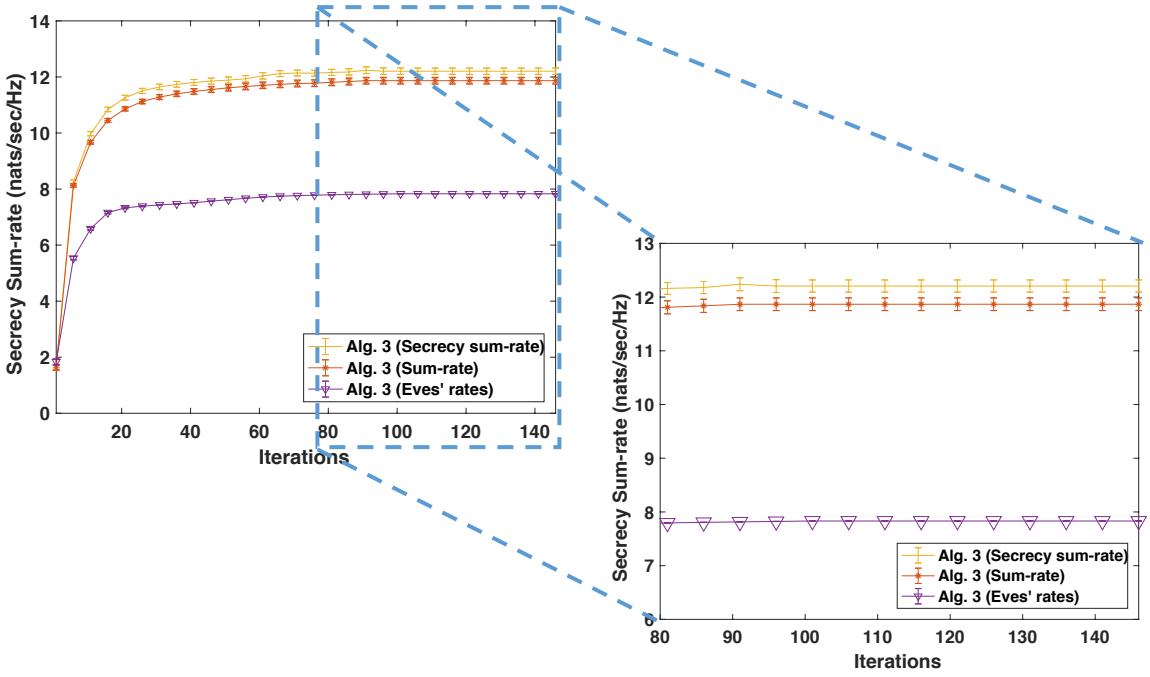
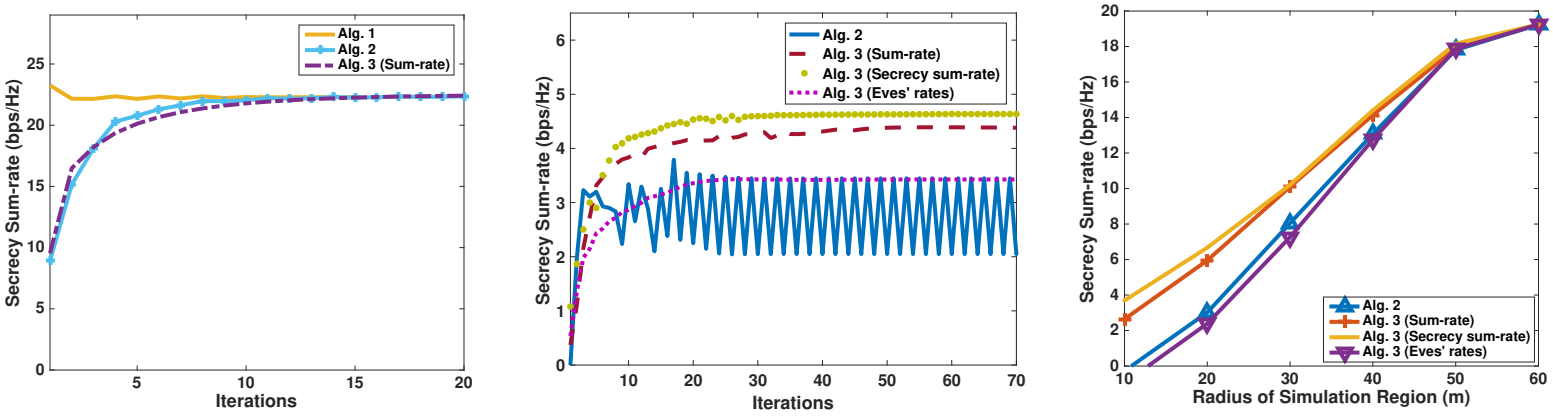


Fig. 4: Comparison of convergence trend of the proposed QNE selection methods: $\big(8$ links $(Q = 8)$ and 7 Eves $(K = 7), r_{circ} = 30$ m, $N_{T_q} = 5, N_{r_q} = 2\ \forall q, N_{e,k} = 2\ \forall k, d_{link} = 10$ m, $P_q = 40$ dBm$\big)$.

100 channel realizations of a given (random) network topology[23]. The results in Fig. 3 (a) and Fig. 3 (b) show that the running time of the QNE selection when secrecy sum-rate is the criterion (i.e., Alg. 3 (Secrecy sum-rate)) is relatively higher than the other two QNE selection methods. It can be seen in Fig. 3 (a) that the difference in the computational complexity of Alg. 3 (Eves' rates) (i.e., QNE selection when minimizing Eves' rates is the criterion) and Alg. 3 (Secrecy sum-rate) appears to be in the slope of the curves, which complies with theoretical derivations in (76) and (77). However, this difference becomes

[23]While we tried to generate the results that are as close as possible to the theoretical derivations, we ended up with non-smooth curves at some points during this simulation. This is mainly due to the fact that in different channel realizations and different initial points, the convergence behavior, and thus the total number of iterations, is not consistent. In order to tackle these problems and generate clean figures, at some points we used linear regression of the actual complexity curve.



(a) (b) (c)

Fig. 5: (a) Convergence of secrecy sum-rate when QNE is unique; (b) convergence of secrecy sum-rate when multiple QNEs exist; (c) secrecy sum-rate vs. r_{circ} : $Q = 8, K = 5, N_{T_q} = 5, N_{r_q} = 2\ \forall q, N_{e,k} = 2\ \forall k, r_{circ} =$ (a) 100 m, (b) 20 m, $P_q =$ (a) 20 dBm, (b) 30 dBm, (c) 40 dBm.

clear when the number of links/antennas are high enough[24]. It can be seen from Fig. 3 (b) that both Alg. 3 (Secrecy sum-rate) and Alg. 3 (Eves' rates) have the same slope. This can be seen in the theoretical derivation for the complexity of both QNE selection methods in (76) and (77), where for both criteria, the complexity w.r.t K is a multiple $QN_{e,k}^3$. For the case of Alg. 3 (Sum-rate) (i.e., QNE selection when maximizing sum-rate is the criterion) the complexity w.r.t K is only a multiple of $N_{e,k}^3$. The gap between the Alg. 3 (Secrecy sum-rate) and Alg. 3 (Eves' rates) in Fig. 3 (b) is because of the additional complexity of Alg. 3 (Secrecy sum-rate), which is independent of the number eavesdroppers (i.e., K).

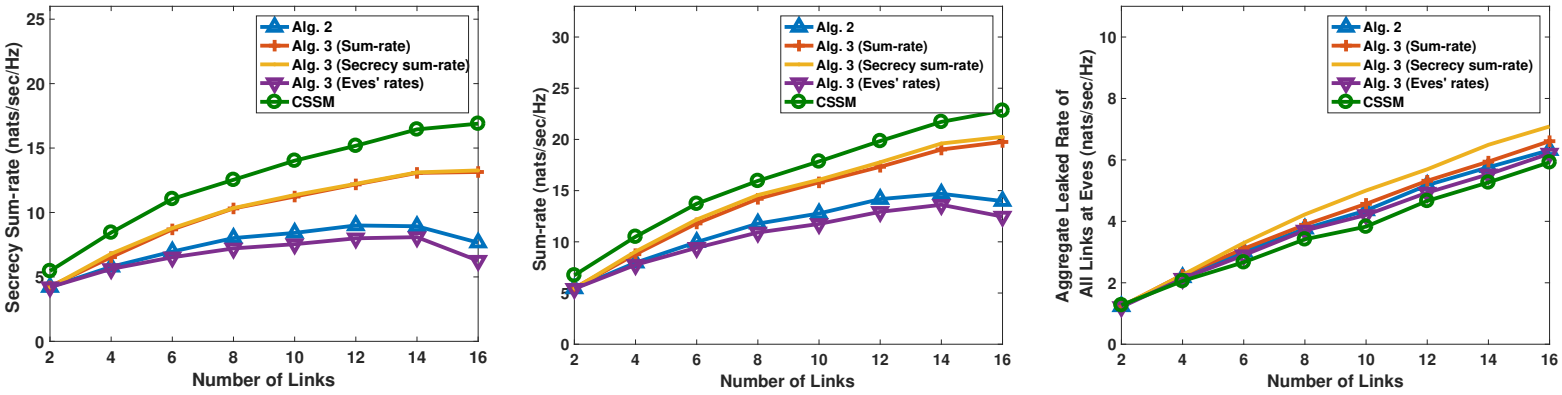
B. Effect of Initial Points

In general, the initial values for the covariance matrices of information and AN signals can affect the results. Given the non-convexity of links' optimization problems, and the fact that at a QNE links operate at their stationary points, which are not necessarily unilaterally optimal, it is theoretically expected that different initial values can make the algorithm converge to different stationary points, thus affecting the final results. However, in our simulations, we did not see any significant variations in the secrecy sum-rate when the initial values of information and AN covariance matrices are changed. For example, by changing the initial values, for networks with 10 to 16 links, a maximum difference of 3 nats/sec/Hz and maximum of 150 iterations until convergence were observed. The results can be seen in Fig. 4, where the simulated convergence behavior of all three QNE selection methods is depicted for one channel realization. A point at the nth iteration of a curve represents the resulting secrecy sum-rate of that particular QNE selection method at the nth iteration, averaged over 100 random initial points. The corresponding 95% confidence intervals are also shown. The tightness of the confidence intervals indicate that while the performance varies when the initial points change, this variation is negligible. Note that in all of our simulations, we considered random initializations for each channel realization of a given (random) network topology.

C. Overall Performance and Energy Efficiency

Fig. 5 (a) compares the three proposed algorithms in a channel realization for the case when the QNE is unique. According to the uniqueness condition in Theorem 3, it is generally expected that if links are far enough from each other, then the resulting QNE is likely to be unique. We simulate this scenario by

[24]Note that the theoretical derivations are derived for the worst case.



(a) (b) (c)

Fig. 6: Comparison of (a) secrecy sum-rate, (b) sum-rate, (c) sum of eavesdroppers' received rates vs. number of links: $r_{circ} = 30$ m, $K = 5$, $N_{T_q} = 5$, $N_{r_q} = 2\ \forall q$, $N_{e,k} = 2\ \forall k$, $d_{link} = 10$ m, $P_q = 40$ dBm.

increasing r_{circ} significantly. We consider the secrecy sum-rate as the measure of comparison between the algorithms. It can be seen that all of the algorithms converge to almost the same point. This result indicates the equivalence between the QNEs found by both Algorithms 1 and 2. Furthermore, it can be concluded that the QNE selection algorithm with sum-rate as its design criterion (indicated by Alg. 3 (Sum-rate)) does not outperform Algorithm 2 when the QNE is unique (i.e., the condition in Theorem 3 is satisfied). That is, if the QNE is unique the QNE selection algorithms only have one QNE to choose from. It should be noted that Algorithm 1 converges faster than other algorithms. This might be because Algorithms 2 and 3 use smaller steps towards the QNE at each iteration.

Fig. 5 (b) compares the achieved secrecy sum-rate in a channel realization between Algorithm 2 and different versions of Algorithm 3, indicated by "Alg. 3 (Secrecy sum-rate)" when secrecy sum-rate is the design criterion, "Alg. 3 (Eves' rates)" when reducing Eves' rates is the design criterion, and "Alg. 3 (Sum-rate)" when sum-rate is the design criterion. Furthermore, due to the existence of multiple QNEs, Algorithm 2 is oscillating between QNEs and never converges even after 70 iterations[25]. We increased the number of iterations to 1000, but did not see the convergence of Algorithm 2. However, all of the versions of Algorithm 3 converge to a QNE[26].

Fig. 5 (c) shows the secrecy sum-rate resulting from different algorithms vs. r_{circ}. For Algorithm 2, we limit the iterations to 100. For Algorithm 3, we limit the iterations of the inner loop (i.e., line 3 in Algorithm 3) and the outer loop (i.e., line 1 in Algorithm 3) to 50 and 3, respectively. Each point in the figure is the result of averaging over 50 random network topologies, where in each topology, 200 channel realizations are simulated and averaged. It can be seen that when r_{circ} is small (i.e., high interference), Alg. 3 (Sum-rate) and Alg. 3 (Secrecy sum-rate) have higher secrecy sum-rate than Algorithm 2. This is due to the fact that the myopic maximization of secrecy rates in Algorithm 2 is not guaranteed to converge to a QNE. Moreover, it can be seen that in Alg. 3 (Eves' rates), we cannot increase the secrecy rate as much as other versions of Algorithm 3. This is due to the fact that in minimizing the received rate at eavesdroppers, too much AN power creates unwanted interference on legitimate receivers, preventing any improvement on the secrecy sum-rate.

Fig. 6 (a) compares the secrecy sum-rate of Algorithms 2 and 3 for different number of links. Alg. 3

[25]Recall that convergence of Algorithm 2 is tied to the uniqueness of the QNE. Furthermore, due to the similarity in the behavior of Algorithms 1 and 2, we only showed Algorithm 2 in subsequent simulations.

[26]The result in Fig. 5 (b) should not be confused with the previous simulation in Fig. 5 (a). In fact, equal secrecy sum-rate for all of the algorithms happen only when QNE is unique (i.e., the condition in Theorem 3 is satisfied). However, Fig. 5 (b) is showing results when the condition in Theorem 3 is not likely to be satisfied.

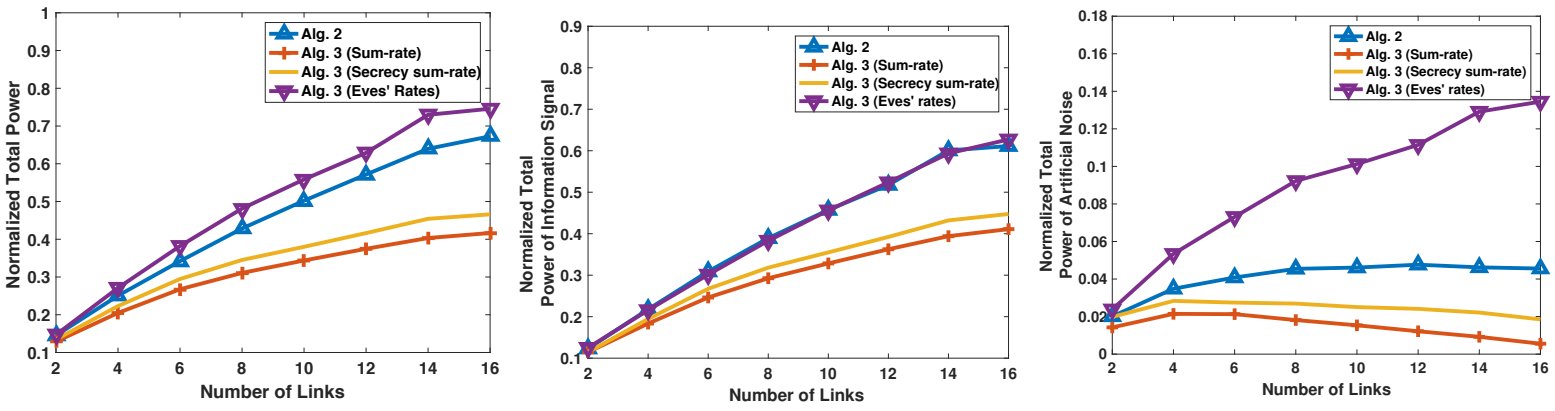


Fig. 7: Comparison of (a) total power (b) total power of information signal (c) total power of AN vs. number of links: $r_{circ} = 30$ m, $K = 5, N_{T_q} = 5, N_{r_q} = 2\ \forall q, N_{e,k} = 2\ \forall k, d_{link} = 10$ m, $P_q = 40$ dBm.

(Secrecy sum-rate) and Alg. 3 (Sum-rate) consistently outperform Algorithm 2 in terms of secrecy sum-rate (Fig. 6 (a)) and sum-rate (Fig. 6 (b)), and Alg. 3 (Eves' rates) does not result in a secrecy sum-rate as high as the other two flavors of Algorithm 3. As shown in Fig. 6 (c), using Alg. 3 (Eves' rates) slightly reduces sum of Eves' received rates by increasing interference at Eves, but this directly affects legitimate transmissions as well. Furthermore, Alg. 3 (Secrecy sum-rate) does not have a significant advantage over Alg. 3 (Sum-rate). Another interesting point is that Alg. 3 (Secrecy sum-rate) has slightly higher sum-rate and higher leaked rate compared to Alg. 3 (Sum-rate). Hence, the performance of Alg. 3 (Secrecy sum-rate) is not necessarily a combination of Alg. 3 (Sum-rate) and Alg. 3 (Eves' rates), but rather a good tradeoff point. Lastly, it can be seen that the proposed algorithms have lower secrecy sum-rates compared to CSSM. We conjecture that this might be due to the fact that CSSM has a larger solution space compared to our methods. Note that the solution space of CSSM may contain some points that are not necessarily the QNEs of the game, whereas both Algorithms 2 and 3 can only converge to QNEs of the game. The difference between Algorithms 2 and 3 is that Algorithm 3 selects the best QNE (according to a criterion), but Algorithm 2 does not. As can be seen in Fig. 6 (a), for the case of 16 links, the loss of Algorithm 3 compared to CSSM is less than 25% when either secrecy sum-rate or sum-rate is the criterion for the QNE selection phase of Algorithm 3. Despite this loss, using Algorithm 3 facilitates not only a distributed implementation, but also the flexibility in the amount of coordination. The latter gives us freedom to keep the coordination as low as possible. Neither of these features are available in CSSM.

In Fig. 7 (a)–(c) the power consumption of different algorithms are compared. The total power in Fig. 7 (a)–(c) is normalized w.r.t the total power budget $\sum_q P_q$. Generally, Alg. 3 (Sum-rate) is the most energy efficient algorithm. Both Alg. 2 and Alg. 3 (Eves' rates) perform poorly in energy efficiency as the increase in the power of AN creates interference at other legitimate receivers. This makes the links to spend even more power on the information signal which eventually leads to neither a high sum-rate nor a high secrecy sum-rate. Moreover, the increase in the power of AN seems to be more significant in Alg. 3 (Eves' rates), as the design criterion forces the users to carelessly increase the interference at Eves. Lastly, Alg. 3 (Secrecy sum-rate) and Alg. 3 (Sum-rate) decrease the power of AN as the number of links increases because as the links abound, they automatically create additional interference at Eves. Hence, the links do not spend more power on AN.

Fig. 8 shows that as the number of eavesdroppers in the network increases, Alg. 3 (Sum-rate) and Alg. 3 (Secrecy sum-rate) outperform Algorithm 2 in terms of secrecy sum-rate, and Alg. 3 (Eves' rates) still achieves a low secrecy sum-rate. Overall, in these simulations, maximizing sum-rate as a design

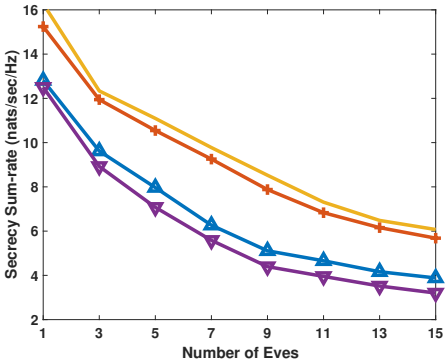


Fig. 8: Comparison of secrecy sum-rate vs. number of Eves: $r_{circ} = 30$ m, $Q = 8, N_{T_q} = 5, N_{r_q} = N_{e,k} = 2, P_q = 40$ dBm

criterion seems to be the best to increase the secrecy sum-rate because other proposed criteria cannot add significant improvements despite requiring more extensive signaling between the links (e.g., the knowledge of all eavesdropping channel gains). Lastly, minimizing Eves' rates as the design criterion although brings poor performance to the QNE selection, it gives us valuable insights on the importance of interference management such that if it is overlooked, the secrecy sum-rate in the network can be severely decreased.

IX. Conclusions

We designed a game theoretic secure transmit optimization for a MIMO interference network with several MIMO-enabled eavesdroppers. We proposed three algorithms to increase secrecy sum-rate. In the first algorithm, the links myopically optimize their transmission until a quasi-Nash equilibrium (QNE) is reached. Because of the inferior performance of first algorithm in case of multiple QNEs, we designed the second algorithm based on the concept of variational inequality. The second algorithm enables us to analytically derive convergence conditions, but achieves the same secrecy sum-rate as the first algorithm. To increase the secrecy sum-rate, we proposed the third algorithm in which the links can select the best QNE according to a certain design criterion. Simulations showed that not every criterion is good for the performance improvement. Specifically, reducing co-channel interference is a better criterion compared to increasing interference at the eavesdroppers to improve secrecy sum-rate.

Appendix A
Proof of Proposition 1

Let $(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K)$ denote the limit point of AO iterations found in Line 10 of Algorithm 1 for the qth link, $q \in \mathbf{Q}$. As mentioned earlier, problem (51) is convex w.r.t either $(\mathbf{\Sigma}_q, \mathbf{W}_q)$ or $\{\mathbf{S}_q\}_{k=0}^K$. Then, recalling the minimum principle in (24), we have the following[27]:

$$X_q = [\mathbf{\Sigma}_q^{*T}, \mathbf{W}_q^{*T}]^T, Z_q = [\mathbf{\Sigma}_q{}^T, \mathbf{W}_q{}^T]^T, \nabla_{Z_q} \bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K) = \left[-(\nabla_{\mathbf{\Sigma}_q} \bar{f}_q)^T, -(\nabla_{\mathbf{W}_q} \bar{f}_q)^T\right]^T, \tag{78a}$$

$$\left\langle Z_q - X_q, \nabla_{Z_q} \bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K) \right\rangle \geq 0, \ \forall (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q, \tag{78b}$$

$$\left\langle \mathbf{S}_{q,k} - \mathbf{S}_{q,k}^*, \nabla_{\mathbf{S}_{q,k}} \bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K) \right\rangle \geq 0, \ \forall \mathbf{S}_{q,k} \succeq 0, \ \forall k \in \mathbf{K}. \tag{78c}$$

It should be noted that for a given $(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*)$, the value of $\{\mathbf{S}_{q,k}^*\}_{k=0}^K$ are uniquely determined (cf. (11b) and (11c)). Hence, using Danskin's theorem [24], the function $\bar{f}_q(\mathbf{\Sigma}_q, \mathbf{W}_q, \{\mathbf{S}_{q,k}^*\}_{k=0}^K)$ is differentiable w.r.t

[27] AO iterations converge to a stationary point of (51) [5, Section IV-B], [25, Corollary 2].

$(\mathbf{\Sigma}_q, \mathbf{W}_q)$, and inequality (78b) holds[28]. Moreover, it can be verified that

$$\nabla_{\mathbf{\Sigma}_q} \bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K) = \nabla_{\mathbf{\Sigma}_q} \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q{}^*), \tag{79}$$

$$\nabla_{\mathbf{W}_q} \bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \{\mathbf{S}_{q,k}^*\}_{k=0}^K) = \nabla_{\mathbf{W}_q} \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q{}^*) \tag{80}$$

where $\bar{R}_{s,q}$ is the smooth approximation of secrecy rate mentioned in (13). Then, according to (79),

$$\left\langle Z_q - X_q, \nabla_Z \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*) \right\rangle \leq 0, \ \forall (\mathbf{\Sigma}_q, \mathbf{W}_q) \in \mathcal{F}_q \tag{81}$$

where $\nabla_Z \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*) = \left[(\nabla_{\mathbf{\Sigma}_q} \bar{R}_{s,q})^T, (\nabla_{\mathbf{W}_q} \bar{R}_{s,q})^T\right]^T$. Hence, $(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*)$ is the optimal solution to

$$\begin{aligned} \underset{Z_q}{\text{maximize}} \quad & \left\langle Z_q - X_q, \nabla_Z \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*) \right\rangle \\ \text{s.t.} \quad & Z_q \in \mathcal{F}_q. \end{aligned} \tag{82}$$

Hence, $(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*)$ must satisfy the K.K.T conditions of (82), which can be written as

$$\nabla_{\mathbf{\Sigma}_q} \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*) - \zeta_q I + \Xi_{q,1} = 0 \tag{83a}$$

$$\nabla_{\mathbf{W}_q} \bar{R}_{s,q}(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*) - \zeta_q I + \Xi_{q,2} = 0 \tag{83b}$$

$$\zeta_q(\text{tr}(\mathbf{\Sigma}_q^* + \mathbf{W}_q^*) - P_q) = 0, \mathbf{\Sigma}_q^* \Xi_{q,1} = 0, \mathbf{W}_q^* \Xi_{q,2} = 0 \tag{83c}$$

$$\zeta_q \geq 0, \Xi_{q,1} \succeq 0, \Xi_{q,2} \succeq 0. \tag{83d}$$

where ζ_q, $\Xi_{q,1}$, and $\Xi_{q,2}$ are Lagrange multipliers. Therefore, the stationary point of AO iterations satisfies the K.K.T conditions of (13).

APPENDIX B
PROOF OF THEOREM 2

To prove the existence of the QNE, we use the following theorem:

Theorem 5. *[11, Corollary 2.2.5] For a mapping $F : \mathcal{Q} \to \mathcal{R}^N$ that is continuous on the compact and convex set $\mathcal{Q} \subseteq \mathcal{R}^N$, the solution set for $VI(F, \mathcal{Q})$ is nonempty and compact.* □

The objective in (51) is continuously differentiable on its domain, making $F^{\mathbb{R}}$ continuous. Furthermore, the set $\mathcal{K}$ is a compact convex set because it is the Cartesian product of compact convex sets (i.e., players' strategy sets). Hence, $\mathcal{K}^{\mathbb{R}}$, the real-vector version of $\mathcal{K}$, is a convex set. Due to the presence of power constraints, the strategy set of each player is compact, then the set $\mathcal{K}^{\mathbb{R}}$ is also compact. Thus, according to Theorem 5, the solution set to the VI in (40) is nonempty, meaning that the QNE in the proposed smooth game exists.

APPENDIX C
PROOF OF THEOREM 3

We first introduce following definition:

Definition 2. *[18, Definition 26] Considering the complex VI in* (25)*, with $F^{\mathbb{C}}(Z) : \mathcal{K} \to \mathbb{C}^{N' \times N}$, $\mathcal{K} \subseteq \mathbb{C}^{N' \times N}$ being a continuously $\mathbb{R}-$differentiable function and $\mathcal{K}$ being a convex set that has a non-empty interior. The augmented Jacobian matrix for $F^{\mathbb{C}}(Z)$, namely, $JF^{\mathbb{C}}(Z)$, is defined as follows[29]:*

$$JF^{\mathbb{C}}(Z) \triangleq \frac{1}{2} \begin{bmatrix} D_Z F^{\mathbb{C}}(Z) & D_{Z^*} F^{\mathbb{C}}(Z) \\ D_Z (F^{\mathbb{C}}(Z)^*) & D_{Z^*} (F^{\mathbb{C}}(Z)^*) \end{bmatrix} \tag{84}$$

[28]Similar reasoning for $\bar{f}_q(\mathbf{\Sigma}_q^*, \mathbf{W}_q^*, \mathbf{S}_{q,k})$ can be used to justify the inequality in (78c).

[29]For the case of $\mathcal{K}$ having a possibly empty interior, the equivalent conidtion in [18, Proposition 28] can be used.

where $D_Z(F^{\mathbb{C}}(Z)) \triangleq \frac{\partial\ vec\left(F^{\mathbb{C}}(Z)\right)}{\partial\ vec(Z)^T}$ *is a* $N'N \times N'N$ *derivative matrix,* $D_{Z^*}F^{\mathbb{C}}(Z)^* = D_Z(F^{\mathbb{C}}(Z)^*$*, and* $D_Z\left(F^{\mathbb{C}}(Z)^*\right) = D_{Z^*}F^{\mathbb{C}}(Z)$*.*

Using this definition, the following proposition holds for $VI(F^{\mathbb{C}}, \mathcal{K})$.

Proposition 3. *[18, Proposition 27] For the* $VI(F^{\mathbb{C}}, \mathcal{K})$ *defined in Definition 1, it holds that:*

- $F^{\mathbb{C}}$ *is monotone on* $\mathcal{K}$ *if and only if* $JF^{\mathbb{C}}(Z)$ *is Augmented Positive Semidefinite (APSD) on* $\mathcal{K}$*. That is, for all* $Y \in \mathbb{C}^{N' \times N}$ *and* $Z \in \mathcal{K}$*,*

$$[vec(Y^*)^T, vec(Y)^T]JF^{\mathbb{C}}(Z)[vec(Y)^T, vec(Y^*)^T]^T \geq 0 \tag{85}$$

  Therefore, $VI(F^{\mathbb{C}}, \mathcal{K})$ *is called a monotone VI and has a (possibly empty) convex solution set.*
- *If* $JF^{\mathbb{C}}(Z)$ *is Augmented Positive Definite (APD) on* $\mathcal{K}$*, then* $F^{\mathbb{C}}$ *is strictly monotone on* $\mathcal{K}$*.* $JF^{\mathbb{C}}(Z)$ *is APD if the inequality in* (85) *is strict. Hence,* $VI(F, \mathcal{Q})$ *is a strictly monotone VI and has at most one solution (if there exists any).*
- $F^{\mathbb{C}}$ *is strongly monotone on* $\mathcal{K}$ *with constant* $c_s > 0$ *if and only if* $JF^{\mathbb{C}}(Z)$ *is uniformly APD on* $\mathcal{K}$ *with constant* $c_s/2$*. That is, for all* $Y \in \mathbb{C}^{N' \times N}$ *and* $Z \in \mathcal{K}$*, there exists a constant* c_s *such that*

$$[vec(Y^*)^T, vec(Y)^T]JF^{\mathbb{C}}(Z)[vec(Y)^T, vec(Y^*)^T]^T \geq c_s||Y||_F^2 \tag{86}$$

  where $||.||_F$ *is the Frobenius norm. Hence,* $VI(F, \mathcal{Q})$ *is a strongly monotone VI and has a unique solution.*

We write the augmented Jacobian matrix for $F^{\mathbb{C}}(\mathbf{\Sigma}, \mathbf{W})$ according to (84). Let $D_Z F^{\mathbb{C}}(Z)$ be defined as

$$D_Z F^{\mathbb{C}}(Z) \triangleq \begin{bmatrix} D_{Z_1}F_1^{\mathbb{C}}(Z_1) & \dots & D_{Z_Q}F_1^{\mathbb{C}}(Z_1) \\ \vdots & \ddots & \vdots \\ D_{Z_1}F_Q^{\mathbb{C}}(Z_Q) & \dots & D_{Z_Q}F_Q^{\mathbb{C}}(Z_Q) \end{bmatrix} \tag{87}$$

where $D_{Z_l}F_q^{\mathbb{C}}(Z_q)$ for all $q, l \in 1, ..., Q^2$ is defined as

$$D_{Z_l}F_q^{\mathbb{C}}(Z_q) \triangleq \begin{bmatrix} D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q) & D_{\mathbf{W}_l}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q) \\ D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{W}_q}\bar{f}_q) & D_{\mathbf{W}_l}(-\nabla_{\mathbf{W}_q}\bar{f}_q) \end{bmatrix}, \tag{88}$$

and $D_{Z^*}F^{\mathbb{C}}(Z) = D_Z(F^{\mathbb{C}}(Z)^* = 0$ (cf. (34)). Thus the matrix $JF^{\mathbb{C}}$ becomes a block diagonal matrix. For a QNE to be unique, the matrix $JF^{\mathbb{C}}$ has to satisfy inequality (85) with strict inequality. Since the game is proved to have at least one QNE (using Theorem 2), and since a strictly monotone VI has at most one solution (if there exists any), then the strict monotonicity of the resulting VI from the game is sufficient to prove the uniqueness of QNE. The strict monotonicity property requires $JF^{\mathbb{C}}$ to be APD. In order to satisfy this condition, we only need $D_Z F^{\mathbb{C}}(Z)$ to be Positive Definite (PD). Given $F^{\mathbb{C}}$ in (34), the entries of $D_{Z_l}F_q^{\mathbb{C}}(Z_q)$ are:

where:

$$D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q) \triangleq \sum_{k=1}^{K}\left(\Lambda_{q,l,k} \otimes \mathbf{G}_{qk}^H\mathbf{S}_{q,k}\mathbf{G}_{qk}\right) - \Psi_{ql}^* \otimes \Psi_{ql}. \tag{89}$$

$$\Psi_{ql} \triangleq -\mathbf{H}_{qq}^H\left(\mathbf{M}_q + \mathbf{H}_{qq}\mathbf{\Sigma}_q\mathbf{H}_{qq}^H\right)^{-1}\mathbf{H}_{ql}, \tag{90}$$

$$\Lambda_{q,l,k} \triangleq \begin{cases} \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\mathbf{G}_{lk}^H\left(\mathbf{S}_{q,k} - \mathbf{M}_{e,q,k}^{-1}\right)\mathbf{G}_{lk} - \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\sum_{k'=1}^{K}\left(e^{\beta\varphi_{e,q,k'}}\mathbf{G}_{lk}^H\left(\mathbf{S}_{q,k} - \mathbf{M}_{e,q,k}^{-1}\right)\mathbf{G}_{lk}\right), & l \neq q, \\ \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\mathbf{G}_{qk}^H\mathbf{S}_{q,k}\mathbf{G}_{qk} - \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\sum_{k'=1}^{K}\left(e^{\beta\varphi_{e,q,k'}}\mathbf{G}_{qk'}^H\mathbf{S}_{q,k'}\mathbf{G}_{qk'}\right), & l = q, \end{cases} \tag{91}$$

and the operator $\otimes$ represents the Kronecker product. Furthermore,

$$D_{\mathbf{W}_l}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q) \triangleq D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{W}_q}\bar{f}_q) = \sum_{k=1}^{K}\left(\Omega_{q,l,k}\otimes\mathbf{G}_{qk}^H\mathbf{S}_{q,k}\mathbf{G}_{qk}\right) - \Psi_{ql}^*\otimes\Psi_{ql} \tag{92}$$

where $\forall(l,q)\in\{1,\ldots,Q\}^2$,

$$\Omega_{q,l,k} \triangleq \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\mathbf{G}_{lk}^H\left(\mathbf{S}_{q,k}-\mathbf{M}_{e,q,k}^{-1}\right)\mathbf{G}_{lk} - \frac{\beta e^{\beta\varphi_{e,q,k}}}{\left(\sum_{k'=1}^{K} e^{\beta\varphi_{e,q,k'}}\right)^2}\sum_{k'=1}^{K}\left(e^{\beta\varphi_{e,q,k'}}\mathbf{G}_{lk'}^H\left(\mathbf{S}_{q,k'}-\mathbf{M}_{e,q,k'}^{-1}\right)\mathbf{G}_{lk'}\right), \tag{93}$$

and the first inequality in (92) holds because both of the derivatives $D_{\mathbf{W}_l}(-\nabla_{\mathbf{\Sigma}_q}\bar{f}_q)$ and $D_{\mathbf{\Sigma}_l}(-\nabla_{\mathbf{W}_q}\bar{f}_q)$ are continuous which implies the symmetry of the Hessian matrix (i.e., equality of mixed derivatives). Lastly,

$$D_{\mathbf{W}_l}(-\nabla_{\mathbf{W}_q}\bar{f}_q) \triangleq \sum_{k=1}^{K}\left(\Omega_{q,l,k}\otimes\mathbf{G}_{qk}^H\mathbf{S}_{q,k}\mathbf{G}_{qk} - \Omega_{q,l,k}\otimes\mathbf{G}_{qk}^H\mathbf{M}_{e,q,k}^{-1}\mathbf{G}_{qk} + \pi_{q,l,k}\otimes\pi_{q,l,k}\right) - \Psi_{ql}^*\otimes\Psi_{ql} \tag{94}$$

where

$$\pi_{q,l,k} \triangleq \mathbf{G}_{qk}^H\mathbf{M}_{e,q,k}^{-1}\mathbf{G}_{lk}. \tag{95}$$

Recalling equations (87) and (88) again, to prove $D_Z F^{\mathbb{C}}(Z)$ is PD, we rely on the generalized Gerschgorin circle theorem [22]. Specifically, for a block matrix $\mathbf{A}$ in which the blocks $A_{ij},\ (i,j)=1,\ldots,M$ are $N\times N$ matrices with complex entries, define the matrix norm $|||\bullet|||$ in $\mathbb{C}^{N\times N}$as follows:

$$|||A_{ij}||| \triangleq \sup_{x\in\mathbb{C}^N}\frac{||A_{ij}x||}{||x||}. \tag{96}$$

where $||\bullet||$ is a vector norm on $\mathbb{C}^N$. Using the Gerschgorin circle theorem, every eigenvalue λ of $\mathbf{A}$ satisfies

$$|||(A_{ii}-\lambda I)^{-1}|||^{-1} \leq \sum_{\substack{k=1\\k\neq i}}^{M}|||A_{i,k}||| \tag{97}$$

for at least one $1\leq i\leq M$, where $|||A^{-1}|||^{-1}\triangleq\inf_{x\in\mathbb{C}^N}\frac{||Ax||}{||x||}$, and I is the identity matrix.

Proposition 4. *[22] If the diagonal block $A_{ii},\ i=1,\ldots,M$ of the block matrix $\mathbf{A}$ are nonsingular and if*

$$|||A_{i,i}^{-1}|||^{-1} \geq \sum_{\substack{k=1\\k\neq i}}^{M}|||A_{i,k}|||,\ i=1,\ldots,M \tag{98}$$

for norm $|||\bullet|||$ in $\mathbb{C}^{N\times N}$ (where $|||A_{i,i}^{-1}|||^{-1}=\inf\limits_{x\in\mathbb{C}^N}\frac{||A_{ij}x||}{||x||}$), then $\mathbf{A}$ is a diagonally dominant matrix. Also if the diagonal blocks are PSD, the condition in (98) *is sufficient for the matrix $\mathbf{A}$ to be PSD.*

We can use the above Gerschgorin circle theorem, Proposition 3, and Proposition 4 on $D_Z F^{\mathbb{C}}(Z)$ defined in (87) to obtain the set of conditions with which the augmented Jacobian matrix $JF^{\mathbb{C}}$ is APSD. We also set the norm $|||\bullet|||$ to be the spectral norm. (i.e., $|||A|||_2=\sqrt{\lambda_{\max}\left(A^HA\right)}$ where $\lambda_{\max}(\bullet)$ denotes the spectral radius of a matrix). Therefore, for $JF^{\mathbb{C}}$ to satisfy the condition in (98), we must have [22,

Chapter 6.1]

$$|\lambda_{q,\min}| \geq \sum_{\substack{q=1 \\ q\neq l}}^{Q} |||D_{Z_l} F_q^{\mathbb{C}}(Z_q)|||_2, \ q = 1, \ldots, Q \tag{99}$$

where $\lambda_{q,\min}$ is the smallest eigenvalue of $D_{Z_q} F_q^{\mathbb{C}}(Z_q)$. Using the strict inequality to (99) –as required by the strict monotonicity– and since the diagonal blocks of $D_Z F^{\mathbb{C}}(Z)$ are already PSD (i.e., $\lambda_{q,\min} \geq 0$ due to concavity of qth player's utility to $(\mathbf{\Sigma}_q, \mathbf{W}_q)$), then the condition in (99) changes to

$$\lambda_{q,\min} > \sum_{\substack{q=1 \\ q\neq l}}^{Q} |||D_{Z_l} F_q^{\mathbb{C}}(Z_q)|||_2, \ q = 1, \ldots, Q \tag{100}$$

REFERENCES

[1] A. D. Wyner, "The wire-tap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Jan. 1975.

[2] S. Goel and R. Negi, "Guaranteeing secrecy using artificial noise," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2180–2189, Jun. 2008.

[3] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas part II: The MIMOME wiretap channel," *IEEE Trans. on Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[4] P.-H. Lin, S.-H. Lai, S.-C. Lin, and H.-J. Su, "On secrecy rate of the generalized artificial-noise assisted secure beamforming for wiretap channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1728–1740, Sep. 2013.

[5] Q. Li, M. Hong, H.-T. Wai, Y.-F. Liu, W.-K. Ma, and Z.-Q. Luo, "Transmit solutions for MIMO wiretap channels using alternating optimization," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1714–1727, Sep. 2013.

[6] O. O. Koyluoglu, H. E. Gamal, L. Lai, and H. V. Poor, "On the secure degrees of freedom in the K-user Gaussian interference channel," in *Proc. IEEE ISIT Conf.*, Jul. 2008, pp. 384–388.

[7] A. Kalantari, S. Maleki, G. Zheng, S. Chatzinotas, and B. Ottersten, "Joint power control in wiretap interference channels," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3810–3823, Jul. 2015.

[8] L. Li, C. Huang, and Z. Chen, "Cooperative secrecy beamforming in wiretap interference channels," *IEEE Signal Process Lett.*, vol. 22, no. 12, pp. 2435–2439, Dec. 2015.

[9] J. B. Rosen, "Existence and uniqueness of equilibrium points for concave N-person games," *Econometrica*, vol. 33, no. 3, pp. 520–534, 1965.

[10] J.-S. Pang and G. Scutari, "Nonconvex games with side constraints," *SIAM J. Optimization*, vol. 21, no. 4, pp. 1491–1522, 2011.

[11] F. Facchinei and J. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*. Springer New York, 2007.

[12] X. Huang, B. Beferull-Lozano, and C. Botella, "Quasi-Nash equilibria for non-convex distributed power allocation games in cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 12, no. 7, pp. 3326–3337, 2013.

[13] G. Scutari and J. S. Pang, "Joint sensing and power allocation in nonconvex cognitive radio games: Nash equilibria and distributed algorithms," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4626–4661, 2013.

[14] A. Goldsmith and S.-G. Chua, "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230, Oct. 1997.

[15] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai, "Compound wiretap channels," *EURASIP J. Wireless Commun. Networks*, no. 5, pp. 1–12, Mar. 2009.

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

[17] J. Nocedal and S. J. Wright, *Numerical Optimization*. Berlin, DE: World Scientific, 2006.

[18] G. Scutari, F. Facchinei, J.-S. Pang, and D. Palomar, "Real and complex monotone communication games," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4197–4231, 2014.

[19] R. T. Rockafellar, "Applications of convex variational analysis to Nash equilibrium," in *Proc. 7th Int. Conference on Nonlinear Anal. and Convex Anal.*, 2011, pp. 173–183.

[20] J. Abadie, *Finite Dimensional Variational Inequalities and Complementarity Problems*. North-Holland, 1967.

[21] G. Scutari, D. Palomar, F. Facchinei, and J.-S. Pang, "Convex optimization, game theory, and variational inequality theory," *IEEE Signal Process Mag.*, vol. 27, no. 3, pp. 35–49, May 2010.

[22] R. A. Horn and C. R. Johnson, Eds., *Matrix Analysis*. New York, NY, USA: Cambridge University Press, 1986.

[23] A. Kannan and U. V. Shanbhag, "Distributed computation of equilibria in monotone nash games via iterative regularization techniques," *SIAM J. Optimization*, vol. 22, no. 4, pp. 1177–1205, 2012.

[24] D. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[25] L. Grippo and M. Sciandrone, "On the convergence of the block nonlinear Gauss-Seidel method under convex constraints," *Operation Research Lett.*, vol. 26, no. 3, pp. 127–136, Apr. 2000.

[26] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Communications Magazine*, vol. 53, no. 6, pp. 21–27, 2015.

[27] A. Mukherjee and A. L. Swindlehurst, "Detecting passive eavesdroppers in the MIMO wiretap channel," in *IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP)*, Mar. 2012, pp. 2809–2812.

[28] C. Shin, R. W. Heath, and E. J. Powers, "Blind channel estimation for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 670–685, 2007.

[29] S. Yatawatta and A. P. Petropulu, "Blind channel estimation in MIMO-OFDM systems with multiuser interference," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1054–1068, 2006.