The Secrecy Capacity of Gaussian MIMO Wiretap Channels Under Interference Constraints

Limeng Dong, Sergey Loyka, and Yong Li

***Abstract*— Secure signaling over multiple-input multiple-output (MIMO) wiretap channel (WTC) is studied under interference and transmit power constraints. The classical MIMO WTC model is extended to interference-limited scenarios, so that interference to other users does not exceed a given threshold while ensuring simultaneously no information leakage to an eavesdropper. The operational secrecy capacity of the Gaussian MIMO WTC under interference and transmit power constraints is rigorously established in two forms (a non-convex max problem and a convex–concave max–min problem), to which per-antenna power constraints can be added as well. Optimal signaling directions are characterized in the general case, from which (tight) upper bounds to the rank of optimal transmit covariance matrix are derived. A sufficient condition for the optimality of beamforming and a necessary condition for optimal full-rank signaling are given. Closed-form rank-1 and high-rank solutions are obtained in the case of zero interference constraints. Sufficient and necessary conditions for non-zero secrecy capacity are established. The results are extended to multi-user scenarios. A sufficient and necessary condition for the unbounded growth of the secrecy capacity with transmit power is obtained. The interplay between transmit and interference power constraints is studied, and its significant impact on optimal signaling is demonstrated (so that neither constraint can be absorbed into the other one in general, as was sometimes suggested in the literature). Overall, these results provide insights into fundamental information-theoretic limits and optimal signaling strategies for secure communications under interference constraints.**

***Index Terms*— MIMO, wiretap channel, secrecy capacity, interference, cognitive radio.**

I. INTRODUCTION

THE ever-growing number of wireless devices and users, explosive demand for higher data rates and quality-of-service are among the key factors of extensive efforts by academia and industry to develop 5G systems [1], [2]. Since the development of such systems are facing a number of challenges, several key technologies have been identified to address them: network densification, massive MIMO, millimeter waves and non-orthogonal multiple access (NOMA) [1], [2]. Some of these key technologies are expected to generate significant interference, which, if not managed properly, may significantly diminish performance gains [3], [4]. In particular, network densification via small cells with aggressive frequency re-use (to improve spectral efficiency) is capable of generating a significant amount of co-channel interference, which is also the case for the NOMA or heterogeneous networks (HetNet) re-using the same spectrum (e.g. cellular/WiFi, macro/femto cells, lisensed/unlisensed etc.) [6], [7]. Additional challenges are posed by device-to-device (D2D), vehicle-to-vehicle (V2V) or machine-to-machine communications (M2M) due to their large numbers as well as decentralized nature of communications [1], [5].

All these factors call for careful interference management (avoidance/cancellation/control) and appropriate signal processing techniques [1]–[4], which is one of the issues addressed in this paper. Since the underlying scenarios, models of interference and related problems in 5G are somewhat similar to those in cognitive radio (CR) systems [8], [9], it is expected that the CR approach will play a role in 5G systems [1], [10]. In particular, the CR paradigm offers new opportunities in overcoming spectrum scarcity by more efficient utilization of the available spectrum in a hybrid (licensed/unlisenced) way [8]. When combined with MIMO techniques, the CR approach allows primary and secondary systems to use the same frequency bands and hence improves significantly the overall spectral efficiency [11]–[13], which is conceptually similar to the approaches to increase spectral efficiency in 5G (e.g. NOMA, HetNet, licensed/unlicesed usage of the spectrum).

Due to the broadcast nature of wireless channels, wireless systems are especially vulnerable to various security threads. This is especially true for unlicensed or hybrid (e.g. CR) systems due to their open architecture and shared use of the same spectrum by many users. A number of possible threads have been identified and studied, including primary user emulation, spectrum sensing data falsification, jamming and eavesdropping [14]. In this respect, physical-layer security approach has emerged as a valuable complement to cryptography-based approaches [15], [16]. In this approach, the secrecy of communications is ensured at the physical layer by exploiting the properties of wireless channels to “hide” transmitted information from eavesdropping. Using this

Manuscript received September 12, 2017; revised January 30, 2018; accepted February 16, 2018. Date of publication April 9, 2018; date of current version July 9, 2018.. The work of L. Dong was supported by the Innovation Foundation for Doctoral Dissertation of Northwestern Polytechnical University under Grant CX201812. This paper was presented in part at 15th Canadian Workshop on Information Theory, Quebec City, Canada, June 11–14, 2017 [33], and in part at the 5th IEEE Global Conference on Signal and Information Processing, Montreal, Canada, November 14–16, 2017 [34].*(Corresponding authors: Limeng Dong; Sergey Loyka.)*

L. Dong was with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada. He is now with the School of Electronics and Information, Northwestern Polytechnical University, Xi’an 710072, China (e-mail: dlm nwpu@hotmail.com).

S. Loyka is with the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada (e-mail: sergey.loyka@ieee.org).

Y. Li is with the School of Electronics and Information, Northwestern Polytechnical University, Xi’an 710072, China (e-mail: ruikel@nwpu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSAC.2018.2824758

approach in combination with multi-antenna (MIMO) systems offers significant opportunities for enhancing the secrecy of wireless communications. The wiretap MIMO channel has emerged as a popular model to establish information-theoretic limits to secure communications [17]–[20]. The key performance metric is the secrecy capacity, defined operationally as the maximum achievable rate subject to reliability (low error probability) and secrecy (low information leakage) criteria [15], which is a counterpart of the regular channel capacity (without the secrecy criterion). The secrecy capacity of the AWGN MIMO wiretap channel (WTC) has been established in [17]–[19] under the full channel state information (CSI) assumption, where in particular the optimality of Gaussian signaling has been shown, and an optimal transmit covariance matrix has been found for a number of special cases in e.g. [21]–[24], while the general case still remains an open problem. An algorithm with provable convergence for global maximization of secrecy rates in the general case was proposed in [25].

In the present paper, we address the two key issues discussed above, interference and secrecy of communications, by extending the classical AWGN MIMO WTC model to interference-constrained scenarios (e.g. CR, NOMA, HetNet) and adding interference constraints so that any feasible signalling must ensure that the interference generated to other users (e.g. primary receivers (PR) in the CR context) does not exceed a ceratin threshold, which we term here "CR MIMO WTC". This significantly changes the problem as the feasible set (of admissible transmit covariance matrices) is *not* isotropic anymore, so that known results (e.g. [17]–[24]) do not apply. It is not even clear whether Gaussian signaling is optimal in such setting (recall that it was far from trivial in [18] and [19] to establish the optimality of Gaussian signalling and some key steps in the proofs exploited the isotropic nature of the feasible set).

The main contributions of this paper are two-fold. First, we establish in Section III the operational secrecy capacity of the Gaussian MIMO wiretap channel under interference constraints in a rigorous way and demonstrate that Gaussian signalling is still optimal, all under the full CSI assumption. We emphasize that the secrecy capacity here is defined operationally as the maximum achievable secrecy rate (subject to reliability and secrecy constraints, in addition to the transmit and interference power constraints, TPC and IPC), rather than formally as the difference of certain mutual information terms without demonstrating their operational significance, as sometimes done in the literature (e.g. in [26]). The operational secrecy capacity of Gaussian CR MISO (i.e. single-antenna receivers) WTC was established earlier in [27] and expressed as a quasi-convex optimization problem, which can be solved as a sequence of convex feasibility problems, and was subsequently re-formulated as a convex problem, for which no closed-form solution is known. The present paper considers the full MIMO case and establishes its secrecy capacity in two alternative forms (a non-convex max problem and a convex-concave max-min problem) under the TPC and IPC, to which per-antenna power constraints (PAC) can be added as well.

Second, we characterise optimal signaling directions in the general case in Section IV and obtain (tight) bounds on the rank of optimal transmit (Tx) covariance matrix, giving sufficient conditions for beamforming (rank-1) to be optimal, which generalizes the respective result in [27] to the full MIMO case, and necessary conditions for full-rank optimal covariance. Motivated by the popularity of linear zero-forcing (ZF) precoding techniques for 5G applications as a balanced approach between performance and complexity [1], [4], we consider zero-IPC constraints in Section V and obtain high-rank and rank-one (beamforming) solutions for optimal Tx covariance matrix, which, when combined together, completely solve the case of 2 Tx antennas. Surprisingly, the characterization of optimal signaling directions under the IPC is the same as *without* the IPC, i.e. optimal signaling is on the positive directions of the difference channel, regardless of what the PR channel is (provided it is of full-rank for non-zero IPC). Contrary to [35] and [36], we do not impose any ad-hoc (sub-optimal) transmission strategies with unknown gap to the capacity, but rather establish the secrecy capacity and characterize optimal (i.e. capacity-achieving) signaling strategies.

Sufficient and necessary conditions for non-zero secrecy capacity are also established in Section IV, which take a different form depending on whether the interference power constraint is zero or not. This partially characterises the scenarios where the physical-layer secrecy approach is feasible under interference constraints. The above results are further extended to a multi-user scenario in Section VI. While the secrecy capacity may saturate as the transmit power increases (with or without interference constraints), a sufficient and necessary condition for the unbounded growth of the secrecy capacity under interference constraint is obtained in Section VII, thus characterizing the scenarios where high secrecy capacity is achievable. Since this condition is simultaneously necessary and sufficient, no more general sufficient condition exists.

The interplay between transmit and interference power constraints is studied in Section VIII and its significant impact on optimal signaling is demonstrated, so that neither constraint can be absorbed into the other one in general, contrary to what was sometimes suggested in the literature [11]. While the secrecy capacity still saturates at high SNR, the IPC affects significantly its behaviour compared to the no-IPC case: it lowers the saturation level, which is attained at much lower SNR, and it makes the saturation hard rather than soft. Contrary to the standard water-filling policy, uniform power allocation is *not* optimal at high SNR under interference constraints, while the low-SNR behaviour is not affected.

Our approach to establish the operational secrecy capacity is based on the method in [17] and [18] and extends it to the interference-constrained settings. In particular, while it is rather straightforward to show that the lower and upper bounds to the secrecy capacity in [18] still hold, it is far more challenging to show that the key saddle-point property in [18] still holds under the interference constraints (which make the feasible set non-isotropic) so that the upper and lower bounds to the secrecy capacity coincide at the saddle point, hence

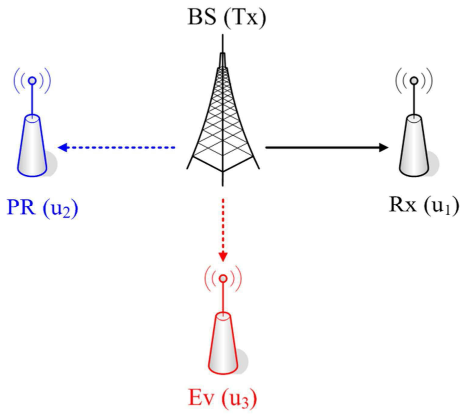

Fig. 1. An example of a network setting where the basestation (BS) is a transmitter that sends a signal to user 1 (Rx) while limiting interference to user 2 (PR) and user 3 (Ev) eavesdrops the transmission.

establishing the operational secrecy capacity. A significant advantage of this approach is that the secrecy capacity is established not only as a non-convex maximization problem but also as a max-min problem which is convex-concave in the right way, so that a globally-optimal numerical algorithm with provable convergence can be developed using the method in [25] and adding interference constraints. This is hardly possible for the non-convex maximization problem (in particular, a provable global convergence is out of reach), for which no closed-form solution is known in the general case either. The established saddle-point in the max-min problem has a game-theoretic interpretation as a minimax game between the transmitter (who selects the input covariance) and nature (who selects the noise covariance); neither player can deviate from the optimal strategy without incurring a penalty.

Notations: bold lower-case letters ($\mathbf{a}$) and capitals ($\mathbf{A}$) denote vectors and matrices respectively; $\mathbf{A}^+$ is Hermitian conjugate of $\mathbf{A}$, $\mathbf{A} \geq \mathbf{0}$ means positive semi-definite; $E\{\cdot\}$ is statistical expectation, $\mathcal{N}(\mathbf{A})$ is the null space of $\mathbf{A}$: $\mathcal{N}(\mathbf{A}) = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}$ while $\mathcal{R}(\mathbf{A})$ denotes the range of $\mathbf{A}$; $\dim\mathcal{N}$ is the dimensionality of $\mathcal{N}$; $\lambda_i(\mathbf{A})$ denotes eigenvalues of $\mathbf{A}$, which are in decreasing order unless indicated otherwise, i.e. $\lambda_1 \geq \lambda_2..$; $r(\mathbf{A})$ denotes the rank of $\mathbf{A}$ while $r_{+(-)}(\mathbf{A})$ is the number of strictly positive (negative) eigenvalues; $|\mathbf{A}|$ and $tr\mathbf{A}$ are determinant and trace; $\mathbf{I}$ is an identity matrix of appropriate size.

II. Channel Model

Let us consider a network setting as in Fig. 1, where base station (BS) is a transmitter (Tx) that sends confidential information to user 1 (a receiver, Rx) while limiting interference to user 2 (PR) and user 3 is an eavesdropper (Ev) that intercepts the transmission. The objective is to ensure reliable communications between the Tx and Rx (the reliability criterion) while keeping the Ev ignorant about transmitted information (the secrecy criterion) and limiting the interference to the PR (user 2) so that its performance is not degraded. The respective wiretap channel model under interference constraint is shown in Fig. 2. For this model, the key performance metric is the secrecy capacity, i.e. the largest transmission rate subject to the reliability and secrecy criteria [15], to which power constraints can also be added.

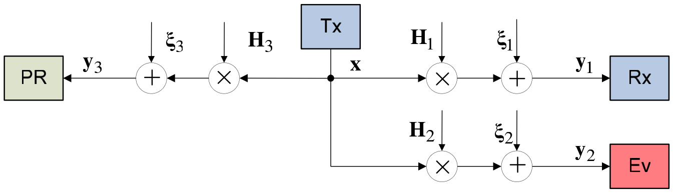

Fig. 2. A block diagram of the Gaussian MIMO wiretap channel under interference constraint. $\mathbf{H}_1$, $\mathbf{H}_2$ and $\mathbf{H}_3$ are the channel matrices to the Rx, Ev and PR respectively; $\mathbf{x}$ is the Tx signal; $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$ are the received signal at the Rx, Ev and PR respectively; $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ and $\boldsymbol{\xi}_3$ are respective noise components.

In the discrete-time channel model, the signals received by the Rx and the Ev are

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2 \tag{1}$$

where $\mathbf{y}_{1(2)}$ are the respective received signals at the Rx and Ev, $\mathbf{x}$ is the transmitted signal, $\boldsymbol{\xi}_{1(2)}$ represent zero-mean unit-variance i.i.d. noise at Rx (Ev) end (so that signal power is also the SNR); $\mathbf{H}_{1(2)}$ are the channel matrices collecting channel gains from all Tx to all Rx(Ev) antennas. In addition to this and following the interference-limited scenario (e.g. cognitive radio, multi-user heterogeneous network or D2D system etc.) there is another user, known as a primary receiver (PR) in the CR setting, and its received signal $\mathbf{y}_3$ is

$$\mathbf{y}_3 = \mathbf{H}_3\mathbf{x} + \boldsymbol{\xi}_3 \tag{2}$$

where $\mathbf{H}_3$ and $\boldsymbol{\xi}_3$ are the channel matrix and zero-mean unit-variance i.i.d. noise of the PR. We assume that the full channel state information (CSI) is available at the Tx, Rx and Ev links; m, n_1, n_2, n_3 are the numbers of antennas at the Tx, Rx, Ev and PR respectively. For future use, let $\mathbf{W}_k = \mathbf{H}_k^+\mathbf{H}_k$, $k = 1, 2, 3$, and note that $\mathbf{W}_k \geq 0$. While the full CSI assumption is justified when an eavesdropper is just another user in the system (and hence shares its CSI with the base station) and is standard in the literature [15]–[27], there are a number of scenarios where this does not hold, in which case one has to consider the compound channel model as in e.g. [37], [38] to account for channel uncertainty. This is beyond the scope of the present paper.

In an interference-limited setting, the transmission is subject to an interference power constraint (IPC), in addition to to the total transmit power constraint (TPC), so that any feasible Tx covariance matrix $\mathbf{R} = E\{\mathbf{x}\mathbf{x}^+\}$ must be in the following feasible set $S_\mathbf{R}$:

$$S_\mathbf{R} = \{\mathbf{R} : tr\mathbf{R} \leq P_T,\ tr\mathbf{W}_3\mathbf{R} \leq P_I,\ \mathbf{R} \geq \mathbf{0}\} \tag{3}$$

where P_T, P_I are the maximum allowed Tx and interference powers respectively. The interference power constraint $tr\mathbf{W}_3\mathbf{R} \leq P_I$ ensures that the total interference power $E\{|\mathbf{H}_3\mathbf{x}|^2\} = tr\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+ = tr\mathbf{W}_3\mathbf{R}$ at the PR (another user) does not exceed the threshold P_I so that its performance is not degraded. The secrecy capacity of the wiretap channel under interference constraint is defined operationally as the largest achievable rate subject to the power, reliability, secrecy, and interference constraints simultaneously.

Without the IPC (i.e. when $\mathbf{W}_3 = 0$), the secrecy capacity C_s of the Gaussian MIMO WTC has been established in [17]–[19]:

$$C_s = \max_{\mathbf{R}\geq 0} C(\mathbf{R}) \quad \text{s.t. } tr\mathbf{R} \leq P_T \tag{4}$$

where $C(\mathbf{R})$ is an achievable secrecy rate ($C(\mathbf{R}) < 0$ is interpreted as zero rate),

$$C(\mathbf{R}) = \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| - \ln|\mathbf{I} + \mathbf{W}_2\mathbf{R}|. \tag{5}$$

It is the purpose of this paper is to extend this result and to establish the secrecy capacity of the (cognitive radio) Gaussian MIMO wiretap channel under interference constraint, which we term here CR MIMO WTC. This task is complicated by the fact that the interference constraint in (3) makes the set $S_{\mathbf{R}}$ non-isotropic in general while the feasible set in (4) is always isotropic and this isotropy was exploited in [18], [19] while establishing the secrecy capacity. In particular, this is critical while establishing a saddle-point and other properties in [18] and subsequently the tight upper and lower bounds which coincide at this saddle point.

III. The Secrecy Capacity of Gaussian CR MIMO WTC

In this section, we establish the operational secrecy capacity of the Gaussian CR MIMO WTC above. To this end, let $S_{\mathbf{K}}$ be a set of covariance matrices of the form

$$S_{\mathbf{K}} \triangleq \left\{\mathbf{K} : \mathbf{K} = \begin{bmatrix} \mathbf{I} & \mathbf{N} \\ \mathbf{N}^+ & \mathbf{I} \end{bmatrix},\ \mathbf{K} \geq \mathbf{0}\right\}, \tag{6}$$

where $\mathbf{N} = E\left\{\boldsymbol{\xi}_1\boldsymbol{\xi}_2^+\right\}$ and

$$f(\mathbf{R},\mathbf{K}) \triangleq \ln|\mathbf{I} + \mathbf{H}^+\mathbf{K}^{-1}\mathbf{H}\mathbf{R}| - \ln|\mathbf{I} + \mathbf{W}_2\mathbf{R}|, \tag{7}$$

where $\mathbf{H} = [\mathbf{H}_1^+, \mathbf{H}_2^+]^+$ is an extended channel. It will be seen later that $\mathbf{K}$ is the covariance matrix of noise $\boldsymbol{\xi} = [\boldsymbol{\xi}_1^+, \boldsymbol{\xi}_2^+]^+$ in the extended channel and $\mathbf{N}$ is the covariance matrix of $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ for an equivalent degraded channel (where these noise vectors are allowed to be correlated with each other).

Theorem 1: The operational secrecy capacity of CR Gaussian MIMO WTC in (1)-(3) *is*

$$C_s = \max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R}) = \max_{\mathbf{R}\in S_{\mathbf{R}}} \min_{\mathbf{K}\in S_{\mathbf{K}}} f(\mathbf{R},\mathbf{K}) \tag{8}$$

Furthermore, the following saddle-point property holds:

$$\max_{\mathbf{R}\in S_{\mathbf{R}}} \min_{\mathbf{K}\in S_{\mathbf{K}}} f(\mathbf{R},\mathbf{K}) = \min_{\mathbf{K}\in S_{\mathbf{K}}} \max_{\mathbf{R}\in S_{\mathbf{R}}} f(\mathbf{R},\mathbf{K}) \tag{9}$$

so that

$$f(\mathbf{R},\mathbf{K}') \leq C_s = f(\mathbf{R}',\mathbf{K}') \leq f(\mathbf{R}',\mathbf{K}) \tag{10}$$

for any feasible $\mathbf{R}$ *and* $\mathbf{K}$*, where* $(\mathbf{R}',\mathbf{K}')$ *is a saddle-point of* $f(\mathbf{R},\mathbf{K})$.

Proof: This result is obtained by establishing tight lower and upper bounds to the secrecy capacity and demonstrating that they coincide provided that $\mathbf{R}$ and $\mathbf{K}$ are selected in an optimal way (i.e. at the saddle point). Since the feasible set $S_{\mathbf{R}}$ is not isotropic[1] anymore (due to the IPC), some key steps of the proofs in [17]–[19] do not hold anymore and new approach is necessary. See Appendix for details. □

A few remarks are in order.

Remark 1: While the first equality in (8) can also be established using the approach of [20], the max-min characterization in (8) as well as the saddle point properties in (9) and (10) cannot be established via that approach. The importance of the max-min characterization comes from the fact that the original max problem in (8) (1st equality) is not convex (unless the channel is degraded) so that all powerful tools of convex optimization cannot be used, while the max-min problem in (8) is convex-concave in the right way[2] and its optimal point is a saddle point. Hence, the Karush-Kuhn-Tucker (KKT) conditions are sufficient for *global* optimality. Furthermore, numerical algorithms with *guaranteed global convergence* can be constructed using the method of [25] by incorporating the extra interference constraints. This is impossible for $\max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R})$ due to its non-convex nature, where global convergence cannot be established, and for which no closed-form solution is known in the general case either.

Remark 2: We emphasize that Theorem 1 characterizes the operational secrecy capacity, defined operationally as the largest achievable secrecy rate, see e.g. [15]- [19]. This is in contrast to other studies where an information capacity is defined formally via the difference of respective mutual information terms without demonstrating their operational significance or where an achievable secrecy rate with unknown gap to the operational capacity is considered, as sometimes done in the literature, see e.g. [26], [35], [36]. The operational characterization of Theorem 1 extends the earlier result established for the CR MISO WTC in [27] to the full MIMO setting.

Remark 3: While $\max\limits_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R})$ and $\max\limits_{\mathbf{R}\in S_{\mathbf{R}}} \min\limits_{\mathbf{K}\in S_{\mathbf{K}}} f(\mathbf{R},\mathbf{K})$ are always the same as (8) indicates, $\mathbf{R}'$ may not be a maximizer of $C(\mathbf{R})$ if the TPC is inactive and $\mathbf{W}_3$ is singular, so that

$$C(\mathbf{R}') < \max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R}) = f(\mathbf{R}',\mathbf{K}') = C_s \tag{11}$$

is possible, in which case an optimal transmit covariance $\mathbf{R}^* = \arg\max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R})$ (i.e. one maximizing the achievable secrecy rate $C(\mathbf{R})$) is not the same as $\mathbf{R}'$ and can only be found from $\max\limits_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R})$; see also an example in Section IX. However,

$$C(\mathbf{R}') = \max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R}) = C_s \tag{12}$$

if either the TPC is active or/and $\mathbf{W}_3$ is non-singular, so that $\mathbf{R}' = \mathbf{R}^*$ in this case. This is also the case without the IPC ($\mathbf{W}_3 = 0$), as in [18], where the TPC is always active.

Remark 4: The saddle-point property in (10) has a well-known game-theoretic interpretation as a minimax game between the transmitter and the nature: neither the transmitter, who controls transmit covariance $\mathbf{R}$, nor the nature, who

[1] A set $S_{\mathbf{R}}$ is isotropic if $\mathbf{R} \in S_{\mathbf{R}}$ implies $\mathbf{URU}^+ \in S_{\mathbf{R}}$ for any unitary $\mathbf{U}$, i.e. eigenvectors of $\mathbf{R}$ are not constrained in any way, only its eigenvalues. This is the case under the TPC only, but not under the IPC in general.

[2] i.e. $f(\mathbf{R},\mathbf{K})$ is concave in $\mathbf{R}$ for any fixed $\mathbf{K}$ and convex in $\mathbf{K}$ for any fixed $\mathbf{R}$.

controls noise covariance $\mathbf{K}$, can deviate from the optimal strategy $(\mathbf{R}', \mathbf{K}')$ without incurring a penalty.

Theorem 1 can be further extended to the case of multiple primary receivers (users) of the form

$$\mathbf{y}_{3k} = \mathbf{H}_{3k}\mathbf{x} + \boldsymbol{\xi}_{3k}, \quad k = 1..K \tag{13}$$

where K is the number of PRs (users) and $\mathbf{H}_{3k}, \ \boldsymbol{\xi}_{3k}$ are the channel matrix and noise for k-th PR. The feasible set of Tx covariance matrices is in this case

$$S'_{\mathbf{R}} = \{\mathbf{R} \geq \mathbf{0} : tr\mathbf{R} \leq P_T, \quad tr\mathbf{W}_{3k}\mathbf{R} \leq P_{Ik}, \ k = 1 \dots K\} \tag{14}$$

where P_{Ik} is the interference constraint power of k-th PR. Theorem 1 applies with $S'_{\mathbf{R}}$ in place of $S_{\mathbf{R}}$.

Using this approach, one can also include per-antenna power constraints (PAC) of the form $r_{ii} \leq P_i$, where r_{ii} is i-th diagonal entry of $\mathbf{R}$ (i.e. power radiated by i-th antenna) and P_i is the i-th antenna constraint power. Specifically, let $\mathbf{W}_{3i}$ be all-zero matrix except for 1 in i-th diagonal entry, so that $tr\mathbf{W}_{3i}\mathbf{R} = r_{ii} \leq P_i$. This can be done in addition to the TPC and IPC.

IV. Characterization of Optimal Signaling Directions

Unfortunately, no closed-form solution is known for the problem in (8) in the general MIMO case. The MISO case (all receivers are equipped with single antennas) was considered in [27], where the respective optimization problem was shown to be quasi-convex, which can be solved numerically as a sequence of convex feasibility problems. It was subsequently transformed into a single convex optimization problem, for which no closed-form solution is known either. The optimal signalling strategy was shown to be beamforming (rank-1).

A closed-form full-rank solution for the problem in (8) under certain conditions was reported in [39] and some unusual properties were discussed, i.e. (i) full-rank solution is not necessarily unique, (ii) strict degradedness ($\mathbf{W}_1 - \mathbf{W}_2 > 0$) is not necessary for a full-rank optimal covariance, and (iii) the TPC can be inactive, which are in stark contrast to the no-IPC case, see e.g. [24].

In this section, we characterize the optimal signaling directions for this problem in the general MIMO case, obtain a (sharp) upper bound for the rank of optimal covariance and establish the optimality of beamforming under certain conditions.

Proposition 1: Let $\mathbf{U}_+$ be a semi-unitary matrix of active eigenvectors (corresponding to strictly positive eigenvalues) of optimal covariance matrix $\mathbf{R}^* = \arg\max_{\mathbf{R} \in S_{\mathbf{R}}} C(\mathbf{R})$ in (8). Then,

$$\mathbf{U}_+^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_+ \geq 0 \tag{15}$$

and the inequality is strict if the TPC is active or/and if $\mathbf{W}_3$ is of full rank. In particular,

$$\mathbf{x}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{x} \geq 0 \tag{16}$$

for any $\mathbf{x} \in \mathcal{R}(\mathbf{U}_+) = \mathcal{R}(\mathbf{R}^*)$ and the inequality is strict under the stated conditions.

Proof: See Appendix. □

Note that, from Proposition 1, the optimal signaling is over non-negative directions of the difference channel $\mathbf{W}_1 - \mathbf{W}_2$ and this holds for any $\mathbf{W}_3$ and any P_I. When $\mathbf{W}_3 > 0$ or/and the TPC is active, this characterization coincides with that in [24] obtained without any interference constraints at all, i.e. the latter has no effect on this characterization (but *does* affect an optimal covariance matrix and the secrecy capacity).

Proposition 1 can be used to establish an upper bound for the rank of optimal covariance $\mathbf{R}^*$. To this end, let $r_{+(-)}(\mathbf{W})$ be the number of positive(negative) eigenvalues of Hermitian matrix $\mathbf{W}$.

Proposition 2: The rank $r(\mathbf{R}^*)$ of optimal covariance $\mathbf{R}^*$ can be upper-bounded as follows:

$$r(\mathbf{R}^*) \leq m - r_-(\mathbf{W}_1 - \mathbf{W}_2) \tag{17}$$

in the general case, and

$$r(\mathbf{R}^*) \leq r_+(\mathbf{W}_1 - \mathbf{W}_2) \leq m - r_-(\mathbf{W}_1 - \mathbf{W}_2) \tag{18}$$

when either the TPC is active or/and $\mathbf{W}_3$ is of full rank.

Proof: See Appendix. □

It is remarkable that these bounds are not affected by $\mathbf{W}_3$ or P_I. Intuitively, this can be understood in the same way as Proposition 1: signaling over negative directions of $\mathbf{W}_1 - \mathbf{W}_2$ provides more information to the Ev and hence cannot ensure secrecy, regardless of the interference constraint, and thus is avoided. The rank bounds follow from the dimensionality of the respective sets. It follows from (18) that

$$r(\mathbf{R}^*) \leq \min(m, n_1),$$

regardless of n_2 and n_3.

Note that 1st inequality in (18) is sharper than (17), since 2nd inequality in (18) can be strict. Indeed, let $r_0(\mathbf{W})$ be the number of zero eigenvalues of $\mathbf{W}$, so that

$$r_+(\mathbf{W}) + r_-(\mathbf{W}) + r_0(\mathbf{W}) = m$$

and hence

$$\begin{aligned} r_+(\mathbf{W}_1 - \mathbf{W}_2) &= m - r_-(\mathbf{W}_1 - \mathbf{W}_2) - r_0(\mathbf{W}_1 - \mathbf{W}_2) \\ &< m - r_-(\mathbf{W}_1 - \mathbf{W}_2) \end{aligned} \tag{19}$$

when $r_0(\mathbf{W}_1 - \mathbf{W}_2) > 0$, i.e. if $\mathbf{W}_1 - \mathbf{W}_2$ is singular.

Optimality of Beamforming: It follows from Proposition 2 that beamforming is optimal, i.e. $r(\mathbf{R}^*) = 1$, if

$$r_-(\mathbf{W}_1 - \mathbf{W}_2) = m - 1 \quad \text{or } r_+(\mathbf{W}_1 - \mathbf{W}_2) = 1$$

and either $\mathbf{W}_3 > 0$ or/and the TPC is active. When the Rx has single antenna, then $r_+(\mathbf{W}_1 - \mathbf{W}_2) = 1$ (unless $C_s = 0$) so that beamforming is optimal, for any number of antennas at the Ev and PR. This extends Corollary 2 in [27] to the case of multi-antenna PR. Note that this rank condition is sufficient but not necessary, i.e. beamforming can sometimes be optimal even when the condition does not hold, e.g. at low SNR.

It also follows from Proposition 2 that $\mathbf{W}_1 > \mathbf{W}_2$, i.e. the channel be strictly degraded, is a necessary condition for the optimal covariance to be of full rank if $\mathbf{W}_3 > 0$ (but it is not sufficient, as can be shown by examples). This does not hold anymore if $\mathbf{W}_3$ is rank-deficient.

Next, we establish sufficient and necessary conditions for non-zero secrecy capacity, thus characterizing the scenarios where the physical-layer security approach is feasible under the interference constraints (see Appendix for a proof).

Proposition 3: The secrecy capacity of the Gaussian MIMO WTC under the interference constraint is non-zero if and only if one of the following holds:

1. $P_I > 0$ and $r_+(\mathbf{W}_1 - \mathbf{W}_2) \geq 1$. Equivalently, $C_s = 0$ iff $\mathbf{W}_1 \leq \mathbf{W}_2$ under non-zero IPC $P_I > 0$.

2. $P_I = 0$, $\mathbf{W}_3$ is rank-deficient, i.e. $|\mathbf{W}_3| = 0$, and $r_+(\mathbf{U}_0^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_0) \geq 1$, where the columns of semi-unitary matrix $\mathbf{U}_0$ form an orthonormal basis of $\mathcal{N}(\mathbf{W}_3)$; they may be taken to be the inactive eigenvectors of $\mathbf{W}_3$ (corresponding to its zero eigenvalues).

It follows from this proposition that $C_s = 0$ if $\mathbf{W}_1 \leq \mathbf{W}_2$, regardless of $\mathbf{W}_3$ and P_I. However, $\mathbf{W}_1 \nleq \mathbf{W}_2$ is not sufficient for $C_s > 0$, unless $P_I > 0$.

V. Optimal Secure Signaling Under Zero Interference Constraint

While Theorem 1 provides a characterization of the secrecy capacity via optimization problems, no closed-form solution is known in the general case. Even in the MISO case, where optimal covariance is known to be rank-1 [27], no closed-form solution is known either, so that one has to resort to numerical algorithms, which limit insights significantly.

To address this issue and motivated by the popularity of linear ZF precoding techniques for 5G applications [1], [4], we consider here the case of zero-IPC, $P_I = 0$, so that the Tx is not allowed to induce any interference power at the PR, and obtain a number of closed-form solutions and related properties. Obviously, this is the most conservative scenario and it provides the largest possible protection to the PR. The respective secrecy capacity will also serve as a lower bound to that of the $P_I > 0$ case.

The following Lemma is needed for further analysis and gives equivalent characterizations of the IPC in this case (see Appendix for a proof).

Lemma 1: The zero interference power constraint $tr\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+ = tr\mathbf{W}_3\mathbf{R} = 0$ can be equivalently expressed as $\mathbf{H}_3\mathbf{R} = \mathbf{0}$ or $\mathbf{W}_3\mathbf{R} = 0$.

It is remarkable that the single (scalar) equality $tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) = 0$ is equivalent to the system of equalities in $\mathbf{H}_3\mathbf{R} = \mathbf{0}$. Note further that $\mathbf{H}_3\mathbf{R} = \mathbf{0}$ represents ZF transmission, which became popular for regular multi-user MIMO systems (no Ev) [28] as well as in the 5G context [1], [4]. It is straightforward to see that $C_s = 0$ unless $\mathbf{W}_3$ is singular, which we assume below.

Using (8) and Lemma 1, the respective secrecy capacity can be expressed as the following problem (P1):

$$(P1): \ C_1 = \max_{\mathbf{R} \in S_{\mathbf{R}}} C(\mathbf{R}) \tag{20}$$

where the feasible set $S_{\mathbf{R}}$ is

$$S_{\mathbf{R}} = \{\mathbf{R} : \mathbf{R} \geq \mathbf{0}, \ tr\mathbf{R} \leq P_T, \ \mathbf{H}_3\mathbf{R} = 0\}. \tag{21}$$

Its operational meaning follows from Theorem 1.

Due to the ZF constraint $\mathbf{H}_3\mathbf{R} = \mathbf{0}$, all columns or eigenvectors of $\mathbf{R}$ are in $\mathcal{N}(\mathbf{H}_3)$. Let $\mathbf{U}_{30}$ be a semi-unitary matrix whose columns are the right singular vectors of $\mathbf{H}_3$ (or, equivalently, eigenvectors of $\mathbf{W}_3$) responsible for zero singular values; they form a basis for $\mathcal{N}(\mathbf{H}_3)$ [32]. Let $\mathbf{P}_3$ be a projection matrix on $\mathcal{N}(\mathbf{H}_3)$, so that

$$\mathbf{P}_3 = \mathbf{U}_{30}(\mathbf{U}_{30}^+\mathbf{U}_{30})^{-1}\mathbf{U}_{30}^+ = \mathbf{U}_{30}\mathbf{U}_{30}^+ \tag{22}$$

from which the following property is obtained.

Lemma 2: Let $\mathbf{R}^*$ be an optimal input covariance of Problem P1. Then,

$$\mathbf{R}^* = \mathbf{P}_3\mathbf{R}^*\mathbf{P}_3 \tag{23}$$

Using this Lemma, we are now in a position to establish an equivalence between the ZF-constrained problem P1 in (20) and an unconstrained but projected problem, where all matrices are projected on the orthogonal basis of $\mathcal{N}(\mathbf{H}_3)$. To this end, let

$$\tilde{\mathbf{W}}_k = \mathbf{U}_{30}^+\mathbf{W}_k\mathbf{U}_{30}, \quad k = 1, 2 \tag{24}$$

and consider the following “projected” problem (P2):

$$(P2): \ C_2 = \max_{\tilde{\mathbf{R}} \in \tilde{S}_{\mathbf{R}}} \ln|\mathbf{I} + \tilde{\mathbf{W}}_1\tilde{\mathbf{R}}| - \ln|\mathbf{I} + \tilde{\mathbf{W}}_2\tilde{\mathbf{R}}| \tag{25}$$

where the feasible $\tilde{S}_{\mathbf{R}}$ set is

$$\tilde{S}_{\mathbf{R}} = \{\tilde{\mathbf{R}} : \tilde{\mathbf{R}} \geq \mathbf{0}, \ tr\tilde{\mathbf{R}} \leq P_T\} \tag{26}$$

Let $r_3 = r(\mathbf{H}_3)$ so that $\dim\mathcal{N}(\mathbf{H}_3) = m - r_3$, and $\tilde{\mathbf{R}}$, $\tilde{\mathbf{W}}_k$ are $(m - r_3) \times (m - r_3)$ matrices. Note that P2 is of lower dimensionality (and hence simpler to solve numerically) than P1 and does not have the ZF constraint anymore. Nevertheless, the following equivalence holds.

Proposition 4: The problems (P1) in (20) *and (P2) in* (25) *are equivalent, so that $C_1 = C_2$ and*

$$\mathbf{R}^* = \mathbf{U}_{30}\tilde{\mathbf{R}}^*\mathbf{U}_{30}^+. \tag{27}$$

where $\mathbf{R}^*$ and $\tilde{\mathbf{R}}^*$ are the corresponding optimal input covariance matrices.

Proof: See Appendix. □

With this equivalence in mind, a number of solutions and results related to the problem P1 can be established. The following Theorem gives a closed-form solution for the optimal covariance matrix when the projected channel is strictly degraded. To this end, let μ_i be i-th eigenvalue of $(\tilde{\mathbf{W}}_2^{-1} - \tilde{\mathbf{W}}_1^{-1})^{-1}$, λ_{min} be the minimum eigenvalue of $\tilde{\mathbf{W}}_1$, and $\alpha = (\mu_1 + \lambda_{min})/\lambda_{min}^2$; $\tilde{\mathbf{U}}$ be a unitary matrix of eigenvectors of $(\tilde{\mathbf{W}}_2^{-1} - \tilde{\mathbf{W}}_1^{-1})$.

Theorem 2: Let the projected channel be strictly degraded, i.e.

$$\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30} > \mathbf{0} \tag{28}$$

and $P_T > P_{T0}$, where P_{T0} is a threshold power given by

$$P_{T0} = 2\alpha \sum_{i=1}^{m-r_3} (\sqrt{1 + 4\mu_i\alpha} + 1)^{-1} - tr(\tilde{\mathbf{W}}_1^{-1}) \tag{29}$$

where $r_3 = r(\mathbf{W}_3)$*. Then,* $r(\mathbf{R}^*) = m - r_3$ *and*

$$\mathbf{R}^* = \mathbf{U}_{30}(\tilde{\mathbf{U}}\tilde{\mathbf{\Lambda}}_1\tilde{\mathbf{U}}^+ - \tilde{\mathbf{W}}_1^{-1})\mathbf{U}_{30}^+ \tag{30}$$

where $\tilde{\mathbf{\Lambda}}_1 = \{\lambda_{1i}\} > 0$ *is a diagonal matrix with*

$$\lambda_{1i} = \frac{2}{\lambda}\left(\sqrt{1+\frac{4\mu_i}{\lambda}}+1\right)^{-1} \tag{31}$$

and $\lambda > 0$ *is found as a unique solution of the following equation:*

$$\sum_{i=1}^{m-r_3} \lambda_{1i} = P_T + \mathrm{tr}(\tilde{\mathbf{W}}_1^{-1}). \tag{32}$$

Proof: See Appendix. □

Theorem 2 gives a closed-form analytical solution for an optimal Tx covariance matrix of high-rank (equal to $m - r_3$, highest possible under zero-IPC) as well as sufficient conditions for this to be the case, i.e. when the SNR exceeds a threshold and the projected channel is strictly degraded. The latter requirement is also necessary for the optimal covariance to be of this high rank, as Proposition 6 below shows. There is no requirement here for the unprojected channel to be degraded (it can be non-degraded, as can be shown by examples). Furthermore, there is no requirement here for SNR → ∞ either, so this is a finite-SNR case. In fact, P_{T0} can be quite small in some scenarios.

It should be pointed out that the solution in Theorem 2 is *not* a projected version of the full-rank solution in [24] without the IPC, i.e. using a no-IPC solution and projecting it orthogonally to $\mathbf{H}_3$ to enforce zero-IPC constraint is *not* optimal. Furthermore, the no-IPC solution in [24] does not apply if the channel is not strictly degraded while the solution in Theorem 2 does apply for a non-degraded channel provided that $\tilde{\mathbf{W}}_1 - \tilde{\mathbf{W}}_2 > 0$.

Using the problem equivalence established in Proposition 4, the following characterization of the optimal signalling directions can be established under zero IPC.

Proposition 5: Let $\mathbf{U}_+$ *be a semi-unitary matrix whose columns are the active eigenvectors* $\{\mathbf{u}_{i+}\}$ *of* $\mathbf{R}^*$*, then*

$$\mathbf{U}_+^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_+ > \mathbf{0} \tag{33}$$

so that $\mathbf{x}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{x} > 0 \ \forall \mathbf{x} \in \mathcal{R}(\mathbf{R}^*) = \mathcal{R}(\mathbf{U}_+)$*, i.e. optimal signalling is on the positive directions of* $\mathbf{W}_1 - \mathbf{W}_2$.

Proof: See Appendix. □

Note that, unlike Proposition 1, there is no requirement here for $\mathbf{W}_3 > 0$ (in fact, $\mathbf{W}_3 > 0$ implies $C_s = 0$ under zero IPC so that $C_s > 0$ requires $\mathbf{W}_3$ to be singular). Furthermore, this characterization is completely independent of the interference constraint (i.e. $\mathbf{W}_3$ or P_I) and coincides with that without the IPC in [24], even though the IPC makes the feasible set here non-isotropic (note also that all matrices in (33) are *unprojected*, as if there were no IPC at all).

Using Proposition 5, the following rank inequality can be established under zero IPC.

Proposition 6: The optimal covariance rank $r(\mathbf{R}^*)$ *can be bounded as follows:*

$$\begin{aligned} r(\mathbf{R}^*) &\le r_+(\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30}) \\ &\le \min\{r_+(\mathbf{W}_1 - \mathbf{W}_2), m - r_3\} \end{aligned}$$

where $r_+(\mathbf{W})$ *is the number of strictly positive eigenvalues of* $\mathbf{W}$.

This Proposition allows one to establish a closed-form beamforming solution for $\mathbf{R}^*$ when

$$r_+(\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30}) = 1,$$

as the next Corollary shows (see Appendix for a proof). A practical importance of this is due to low-complexity mobile units using single antennas (so that the channel rank is automatically 1).

Corollary 1: Let $r_+(\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30}) = 1$*, then* $r(\mathbf{R}^*) = 1$ *for the problem P1 in* (20)*, and*

$$\mathbf{R}^* = P_T\mathbf{U}_{30}\mathbf{u}_1\mathbf{u}_1^+\mathbf{U}_{30}^+ \tag{34}$$

where $\mathbf{u}_1$ *is the eigenvector corresponding to the largest eigenvalue of* $(\mathbf{I}+P_T\tilde{\mathbf{W}}_2)^{-1}(\mathbf{I}+P_T\tilde{\mathbf{W}}_1)$*, so that beamforming along* $\mathbf{U}_{30}\mathbf{u}_1$ *is optimal.*

Note that Theorem 2 and Corollary 1, when combined together, completely solve the case of $m - r_3 \le 2$ (e.g. if $m \le 2$), where either $r(\mathbf{R}^*) = 1$, so that Corollary 1 applies, or $\tilde{\mathbf{R}}^*$ is full-rank, so that Theorem 2 applies. The general case with $m - r_3 > 2$ remains an open problem. Similarly to the high-rank solution in Theorem 2, the rank-1 solution in (34) is *not* a projected version of the no-IPC rank-1 solution (see e.g. [17], [23]) and the former applies even if $r_+(\mathbf{W}_1 - \mathbf{W}_2) > 1$ (unlike the latter) provided that $r_+(\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30}) = 1$.

Next, we consider the case of $\mathcal{R}(\mathbf{W}_2) \in \mathcal{R}(\mathbf{W}_3)$, i.e. when the PR can "see" in all the directions where the Ev can.

Proposition 7: Let $\mathcal{R}(\mathbf{W}_2) \in \mathcal{R}(\mathbf{W}_3)$*. Then, the optimal covariance is*

$$\mathbf{R}^* = \mathbf{U}_{30}(\mu^{-1}\mathbf{I} - \tilde{\mathbf{W}}_1^{-1})_+\mathbf{U}_{30}^+ \tag{35}$$

where $\mu > 0$ *is found from the total power constraint* $tr\mathbf{R}^* = tr(\mu^{-1}\mathbf{I} - \tilde{\mathbf{W}}_1^{-1})_+ = P_T$ *and* $(\mathbf{A})_+$ *denotes positive eigenmodes (corresponding to strictly positive eigenvalues) of Hermitian matrix* $\mathbf{A}$*. The secrecy capacity is*

$$C_s = C(\mathbf{R}^*) = \sum_{i:\lambda_i(\tilde{\mathbf{W}}_1)>\mu} \ln(\lambda_i(\tilde{\mathbf{W}}_1)\mu^{-1}) \tag{36}$$

Proof: See Appendix. □

Finally, we note that these results can also be extended to the case of multiple PRs (users), for which there are multiple zero-IPCs of the form $tr\mathbf{W}_{3k}\mathbf{R} = 0$, $k = 1..K$, equivalent to $\mathbf{H}_{3k}\mathbf{R} = 0$ (as established in Lemma 1), by aggregating their respective channel matrices $\mathbf{H}_{3k}$ into the single matrix $\mathbf{H}_3$,

$$\mathbf{H}_3 = [\mathbf{H}_{31}^+, \mathbf{H}_{32}^+, \ldots, \mathbf{H}_{3K}^+]^+ \tag{37}$$

and using the results above with the new $\mathbf{H}_3$ or equivalently with $\mathbf{W}_3 = \sum_k \mathbf{W}_{3k}$.

VI. An Extension to Multi-User Scenarios

The results of the previous sections can also be extended to some multi-user scenarios, as in Fig. 3, where there are multiple Evs and PRs with respective channel matrices $\mathbf{H}_{21}, \ldots, \mathbf{H}_{2N}$ and $\mathbf{H}_{31}, \ldots, \mathbf{H}_{2K}$, where N and K are the numbers of Evs and PRs respectively. In this scenario,

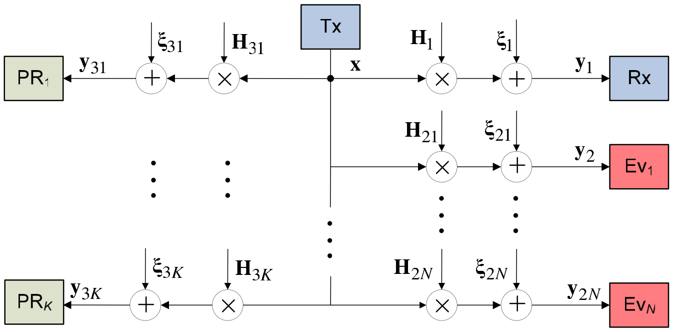

Fig. 3. A block diagram of multi-user Gaussian MIMO wiretap channel under interference constraints. $\mathbf{H}_1$, $\mathbf{H}_{2k}$ and $\mathbf{H}_{3k}$ are the channel matrices to the Rx, k-th Ev and PR respectively. Secrecy and interference constraints are to be satisfied for each Ev and PR respectively.

the secrecy criterion has to be satisfied for each Ev and the IPC - for each PR, i.e. we consider the case of non-cooperating Evs and non-cooperating PRs.

Proposition 8: Consider the non-cooperative multi-Ev multi-PR scenario as in Fig. 3. Assume that there are a dominant Ev and a dominant PR, i.e.

$$\mathbf{W}_{21} \geq \mathbf{W}_{2k}, \quad \mathbf{W}_{31} \geq \mathbf{W}_{3k}, \ \forall k \tag{38}$$

Then, the secrecy capacity under the interference constraints $tr\mathbf{W}_{3k}\mathbf{R} \leq P_I \ \forall k$ *is as in Theorem 1 with* $\mathbf{W}_2 = \mathbf{W}_{21}, \ \mathbf{W}_3 = \mathbf{W}_{31}$.

Proof: See Appendix. □

It follows from this Proposition that if there are a dominant Ev and a dominant PR in the multi-user setting, then a single-user solution (optimal signaling or wiretap code) also works in the multi-user setting, which is a practically-useful robustness property. The dominance condition in (38) essentially means that 1st Ev and 1st PR enjoy the most favorable propagation conditions (e.g. are located closer to the Tx while all others are further away).

The case of cooperating Evs can also be considered in a similar way. Specifically, let $\mathbf{H}_{21} \ldots \mathbf{H}_{2N}$ represent the respective Ev channels and aggregate all of them into the single matrix: $\mathbf{H}_2 = [\mathbf{H}_{21}^+, \ldots, \mathbf{H}_{2N}^+]^+$. Then, Proposition 8 applies with the single aggregate Ev $\mathbf{W}_2 = \mathbf{H}_2^+\mathbf{H}_2 = \sum_k \mathbf{W}_{2k}$ under the second condition in (38). It is straightforward to see that the secrecy capacity for the cooperating Evs never exceeds that for the non-cooperating one, so that this can serve as a lower bound to the secrecy capacity in the latter case (no need for a dominant Ev to exist).

VII. On Unbounded Growth of the Secrecy Capacity

While Proposition 3 above characterises the scenarios where the physical-layer security approach is feasible under interference constraints, it does not tell us whether high secrecy capacity is achievable, which is important for 5G systems. It is well-known that the capacity of the regular Gaussian MIMO channel (no Ev, no PR) grows unbounded as the Tx power increases (even though the growth slows down at high SNR), so that any high capacity can be attained given enough power budget. The behavior changes dramatically for the WTC: the secrecy capacity of Gaussian MIMO WTC may saturate, i.e. stay bounded, $C_s < \infty$, even if $P_T \to \infty$: for example,

$$C_s \to \ln \frac{|\mathbf{W}_1|}{|\mathbf{W}_2|} < \infty \quad \text{as } P_T \to \infty \tag{39}$$

if $\mathbf{W}_1 \geq \mathbf{W}_2 > 0$ [18], [24], see Fig. 7, – a dramatic difference to the regular MIMO channel, so that arbitrary large capacity cannot be attained, even with unlimited power budget. Sufficient conditions for unbounded growth of the secrecy capacity in Gaussian MIMO WTC have been given in [22] for $m = 2$ or when the optimal covariance is of rank 1. In the following, we consider the general case (not limited to $m = 2$ or $r(\mathbf{R}^*) = 1$) and give sufficient and necessary condition for the unbounded growth of the secrecy capacity, which can be further extended to the interference-constrained setting.

Proposition 9: Consider the Gaussian MIMO WTC (no PR) under the Tx power constraint. Its secrecy capacity grows unbounded as the power increases, i.e. $C_s \to \infty$ *as* $P_T \to \infty$, *if and only if*

$$\mathcal{N}(\mathbf{W}_2) \notin \mathcal{N}(\mathbf{W}_1) \tag{40}$$

Proof: see Appendix. □

Note that the condition in (40) cannot be satisfied if $\mathbf{W}_2 > 0$, since $\mathcal{N}(\mathbf{W}_2) = \varnothing$ in this case, so that $\mathbf{W}_2 > 0$ ensures that the secrecy capacity saturates. All other special cases considered in [22] can also be derived from (40). No more general condition for the unbounded growth of the secrecy capacity exists since (40) is also necessary.

The following sufficient condition follows directly from (40).

Corollary 2: A sufficient condition for the unbounded growth of the secrecy capacity of the Gaussian MIMO WTC is that $r(\mathbf{W}_1) > r(\mathbf{W}_2)$.

Proof: Note that $r(\mathbf{W}_k) = m - \dim\mathcal{N}(\mathbf{W}_k)$ so that $r(\mathbf{W}_1) > r(\mathbf{W}_2)$ implies $\dim\mathcal{N}(\mathbf{W}_1) < \dim\mathcal{N}(\mathbf{W}_2)$ and hence (40). □

While this last condition is sufficient, it is not necessary (this can be shown by examples).

As the above results show, the secrecy capacity of the Gaussian MIMO WTC may grow unbounded. One may wonder whether this is still the case under the interference constraint and, if so, under what conditions. The following establishes sufficient and necessary condition for the unbounded growth of the secrecy capacity for the CR MIMO WTC.

Proposition 10: Consider the Gaussian CR MIMO WTC under the TPC and IPC. Its secrecy capacity grows unbounded as the Tx power P_T *increases and* P_I *stays bounded if and only if*

$$\mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_3) \notin \mathcal{N}(\mathbf{W}_1) \tag{41}$$

On the other hand, if $P_T, P_I \to \infty$ *simultaneously, then* (40) *is both sufficient and necessary for* $C_s \to \infty$.

Proof: see Appendix. □

It can be further shown that, under the condition in (41), the secrecy capacity can be lower bounded as follows:

$$\begin{aligned} C_s &\geq \max_{\mathbf{U}} \ln|\mathbf{I}+P_T\mathbf{U}^+\mathbf{W}_1\mathbf{U}/r(\mathbf{U})| \\ &\geq \sum_{i=0}^{d-1} \ln(1+P_T\lambda_{r_1-i}(\mathbf{W}_1)/d) \end{aligned} \tag{42}$$

where $\mathbf{U}$ is any semi-unitary matrix whose columns are a part of an orthonormal basis of $\mathcal{S}$,

$$\mathcal{S} = \mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_3) \cap \mathcal{R}(\mathbf{W}_1) \tag{43}$$

$d = \dim \mathcal{S} \geq r(\mathbf{U})$, $r_1 = r(\mathbf{W}_1)$ and hence $\lambda_{r_1}(\mathbf{W}_1) > 0$ is the minimum non-zero eigenvalue of $\mathbf{W}_1$. The lower bound is achieved by a ZF transmission of the form

$$\mathbf{R} = d^{-1}P_T\mathbf{U}_d\mathbf{U}_d^+ \tag{44}$$

where semi-unitary matrix $\mathbf{U}_d = [\mathbf{u}_{r_1-d+1}, .., \mathbf{u}_{r_1}]$ collects the eigenvectors of $\mathbf{W}_1$ corresponding to its d smallest non-zero eigenvalues. This is also an isotropic signaling over the subspace $\mathcal{S}$.

Equivalently, the sufficient and necessary condition for the secrecy capacity to saturate is that

$$\mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_3) \in \mathcal{N}(\mathbf{W}_1) \tag{45}$$

which automatically holds if $\mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_3) = \varnothing$, e.g. if either $\mathbf{W}_2$ or $\mathbf{W}_3$ are full-rank.

These results can also be extended to the case of multi-PR scenario, where $\mathbf{W}_{31} \ldots \mathbf{W}_{3K}$ represent multiple PRs: (41) extends to

$$\mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_{31}) .. \cap \mathcal{N}(\mathbf{W}_{3K}) \notin \mathcal{N}(\mathbf{W}_1) \tag{46}$$

Equivalently, one can define $\mathbf{W}_3 = [\mathbf{W}_{31}^+, \ldots, \mathbf{W}_{3K}^+]^+$ and use (41) with the new (aggregated) $\mathbf{W}_3$. We remark that, under this condition, unbounded growth of the secrecy capacity can be achieved with ZF transmission.

VIII. On the Activity of Constraints

In the scalar case ($m = 1$), one constraint always dominates and hence only one constraint is active, unless the thresholds are precisely adjusted so that both constraints become identical (a trivial case). A question arises as to whether this is still the case with $m > 1$, or two constraints can be simultaneously active in a non-trivial way? The following example gives a positive answer to this question:

$$\mathbf{W}_1 = \mathbf{I}, \quad \mathbf{W}_2 = 0.5\mathbf{I}, \quad \mathbf{W}_3 = \mathrm{diag}\{1, 2\} \tag{47}$$

It can be shown that the optimal covariance is also diagonal in this case, $\mathbf{R}^* = \mathrm{diag}\{p_1, p_2\}$. Fig. 4 shows the $P_T - P_I$ plane and the regions where the TPC, IPC or both are active. Clearly, both constraints can be active in a non-trivial way and at least one constraint is always active. The respective regions and power allocations are:

Region 1: only the TPC is active if $P_I > 1.5P_T$, and $p_1 = p_2 = P_T/2$, i.e. isotropic signaling is optimal.

Region 2: both the TPC and the IPC are active simultaneously if

$$\max(P_T, f(P_T)) \leq P_I \leq 1.5P_T,$$

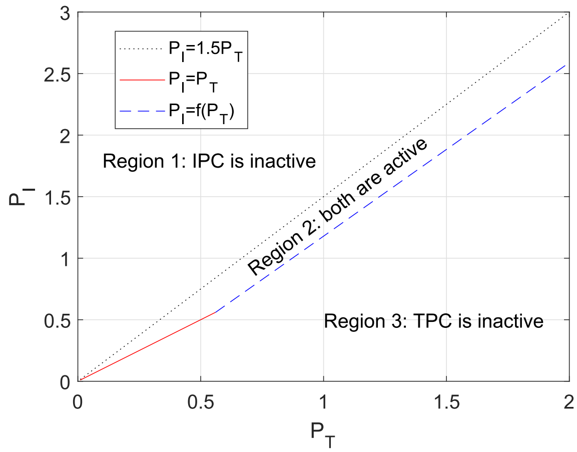

Fig. 4. $P_T - P_I$ plane and the activity regions of the TPC and the IPC for the channel in (47). Both constraint can be active simultaneously in a non-trivial way as in Region 2.

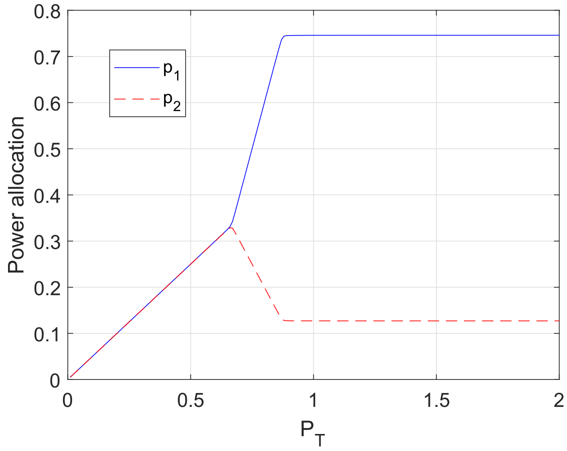

Fig. 5. Optimal power allocation (p_1, p_2) versus P_T for $P_I = 1$. Only the TPC is active if $P_T < 0.67$. Both the TPC and the IPC are active if $0.67 < P_T < 0.87$. Only the IPC is active when $P_T > 0.87$. Different regions correspond to significantly different power allocations. In particular, uniform power allocation is optimal only if $P_T < 0.67$ and sub-optimal otherwise - stark difference to the classical water-filling algorithm, where uniform power allocation becomes optimal at high SNR.

where $f(x) = \sqrt{2(x+3)^2 + 0.25} - 4.5$, and $p_1 = 2P_T - P_I$, $p_2 = P_I - P_T$. Note that optimal signaling is not isotropic anymore and both constraints are active simultaneously in a non-trivial way (in the scalar case of $m = 1$ they would be active simultaneously only if $P_T = aP_I$, i.e. over a line rather than a region as in Fig. 4).

Region 3: only the IPC is active if $P_I < \max(P_T, f(P_T))$ so that $p_1 = P_I$, $p_2 = 0$ if $P_T \geq f(P_T)$, or

$$\begin{aligned} p_1 &= (\sqrt{1 + 4/g^{-1}(P_I)} - 3)/2, \\ p_2 &= (\sqrt{1 + 2/g^{-1}(P_I)} - 3)/2 \end{aligned} \tag{48}$$

if $f(P_T) \geq P_T$, where $g^{-1}(\cdot)$ is the inverse function of

$$g(x) = (\sqrt{1 + 4/x} + 2\sqrt{1 + 2/x} - 9)/2. \tag{49}$$

Fig. 5 and 6 show optimal power allocation versus P_T and P_I respectively for this channel. Clearly, the presence of the extra constraint makes the behavior of the optimal power

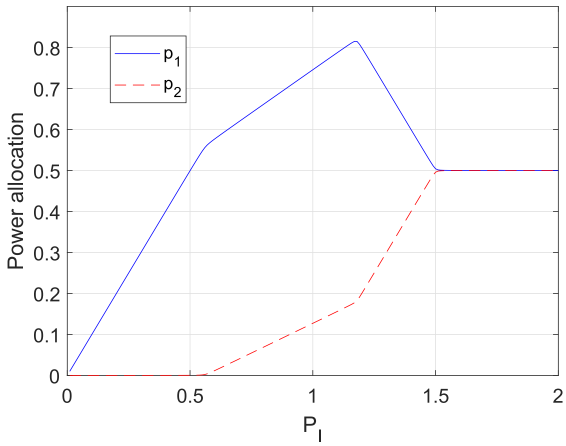

Fig. 6. Optimal power allocation (p_1, p_2) versus P_I for $P_T = 1$. Only the IPC is active if $P_I < 1.18$. Both the TPC and the IPC are active if $1.18 < P_I < 1.5$. Only the TPC is active if $P_I > 1.5$, in which case optimal signaling is isotropic.

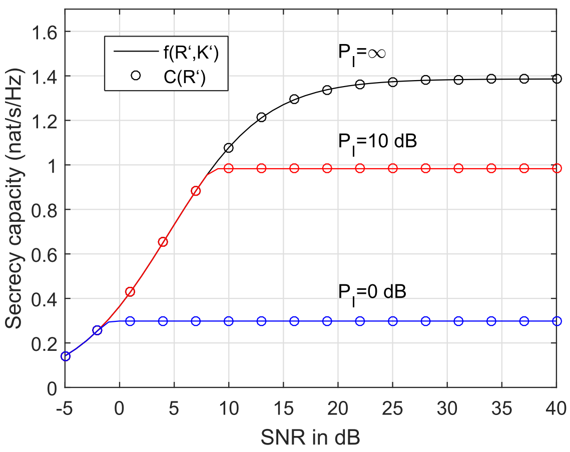

Fig. 7. Secrecy capacity of the channel in (47) vs. the SNR. Note significant loss in the capacity due to the IPC and that the upper $f(\mathbf{R}, \mathbf{K})$ and lower $C(\mathbf{R})$ bounds coincide at the saddle-point $(\mathbf{R}', \mathbf{K}')$ (found via extensive Monte-Carlo simulations), as established in Theorem 1 (see also Remark 3).

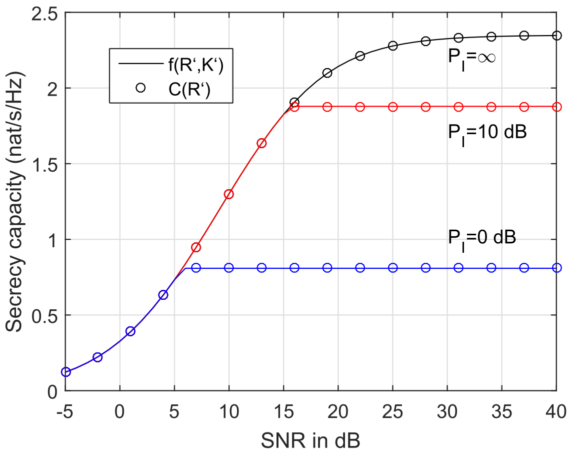

Fig. 8. Secrecy capacity of the channel in (50) vs. the SNR. Note the same tendencies as in Fig. 7. While the high-SNR behaviour is significantly affected by the IPC, there is no change in the low-SNR behaviour.

allocation more complicated, with different regions behaving is a different way, depending on which constraint is active. Note that, in Fig. 5, uniform power allocation is optimal only if $P_T < 0.67$ (inactive IPC) and sub-optimal otherwise - a stark difference to the classical water-filling algorithm, where the uniform power allocation becomes optimal at high power (SNR). This tendency is also observed in Fig. 6, where the uniform power allocation (which corresponds to isotropic signaling) is optimal only if $P_I \geq 1.5$ (inactive IPC) and sub-optimal otherwise.

Fig. 7 illustrates the secrecy capacity dependence on the SNR ($= P_T$) for the channel in (47). Note that there is a significant capacity loss due to the IPC as compared to the no-IPC case (IPC $= \infty$) at high SNR ($> -2(8)$ dB for $P_I = 0(10)$ dB respectively) and that the capacity saturates in this regime in both cases. However, there is a significant difference: in the IPC case, the saturation starts earlier, it is at a lower level and it is hard rather than soft in the no-IPC case: no increase at all after about -2 dB for $P_I = 0$ dB and after 8 dB for $P_I = 10$ dB, which is due to the inactive TPC and this is impossible in the no-IPC case, where the increase becomes smaller with SNR, but never stops completely, approaching the limit at significantly higher SNR of about 25 dB.

IX. Examples

Extensive numerical experiments have been carried out to validate the analytical results above and no difference had been found between the theory and the simulations. To further illustrate the analytical results, we consider the following representative example with the channel matrices $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ set as

$$\begin{pmatrix} 0.2 & 0.1 \\ 0.1 & 1.2 \end{pmatrix}, \quad \begin{pmatrix} 0.5 & 0.9 \\ 0.9 & 1.8 \end{pmatrix}, \quad \begin{pmatrix} 0.6 & 0.5 \\ 0.5 & 1.0 \end{pmatrix} \tag{50}$$

respectively for which the results are shown in Fig. 8. Note that this channel is not degraded (the eigenvalues of $\mathbf{W}_1 - \mathbf{W}_2$ are 0.4 and -1.3) and that the general tendencies are the same as in Fig. 7, even though here channel matrices are not diagonal. Similar tendencies were observed for a large number of cases in our experiments.

To illustrate that $C(\mathbf{R}') < f(\mathbf{R}', \mathbf{K}')$ is possible at saddle-point for singular $\mathbf{W}_3$ and hence $\mathbf{R}^* \neq \mathbf{R}'$, we consider the following example:

$$\mathbf{H}_1 = \text{diag}\{1, 0\}, \quad \mathbf{H}_2 = \text{diag}\{0, 1\}, \quad \mathbf{W}_3 = \text{diag}\{1, 0\} \tag{51}$$

as illustrated in Fig. 9 for $P_I = 1$. It is straightforward to see that the saddle-point is

$$\mathbf{K}' = \mathbf{I}, \quad \mathbf{R}' = \text{diag}\{\min(P_T, P_I), a\} \tag{52}$$

where $0 \leq a \leq (P_T - P_I)_+$, i.e. $\mathbf{R}'$ is not unique if $P_T > P_I$, and the optimal covariance $\mathbf{R}^* = \arg\max_{\mathbf{R} \in S_\mathbf{R}} C(\mathbf{R})$ is

$$\mathbf{R}^* = \text{diag}\{\min(P_T, P_I), 0\} \tag{53}$$

Thus, $\mathbf{R}^* \neq \mathbf{R}'$ (unless $a = 0$) and

$$C_s = f(\mathbf{R}', \mathbf{K}') = \ln(1 + \min(P_T, P_I)) \tag{54}$$

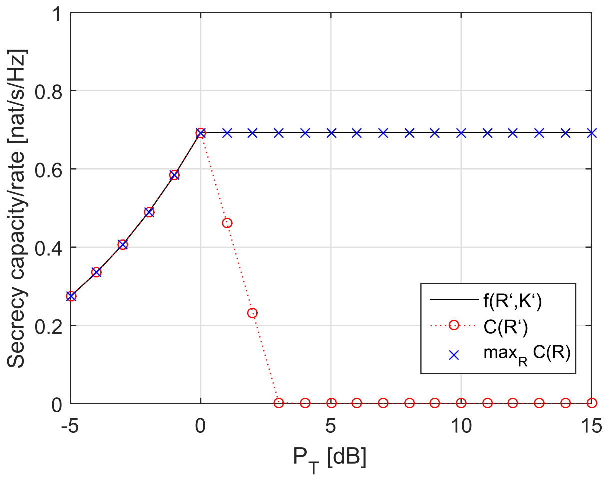

Fig. 9. Secrecy capacity/rate of the channel in (51) vs. P_T; $P_I = 1$. Note that $C(\mathbf{R}') < f(\mathbf{R}', \mathbf{K}')$ when $P_T > P_I$ while $f(\mathbf{R}', \mathbf{K}') = \max_{\mathbf{R}} C(\mathbf{R}) = C(\mathbf{R}^*)$ for any P_T, so that $\mathbf{R}' \neq \mathbf{R}^*$ is possible, unlike that in [18].

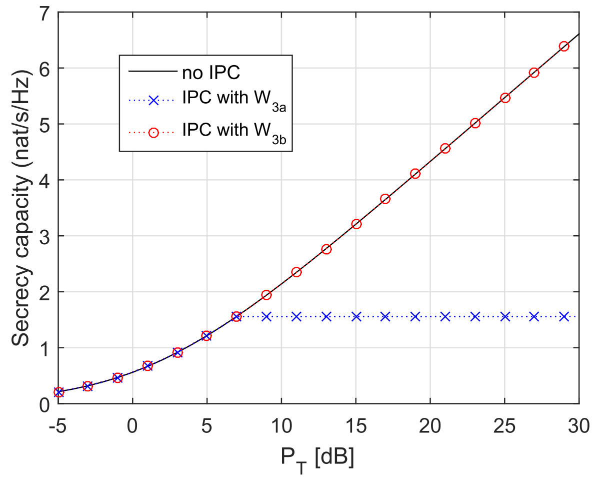

Fig. 10. Unbounded growth of the secrecy capacity for the channel in (56) with P_T; $P_I = 1$. While $\mathbf{W}_{3a}$ bounds the capacity growth, $\mathbf{W}_{3b}$ does not, as expected from Proposition 10.

for any a. However, if one sets $a = (P_T - P_I)_+$, then

$$C(\mathbf{R}') = \ln \frac{1 + \min(P_T, P_I)}{1 + (P_T - P_I)_+} < C_s = C(\mathbf{R}^*) \quad (55)$$

where the inequality holds if $P_T > P_I$, as Fig. 9 shows (negative $C(\mathbf{R}')$ is interpreted as zero rate). Hence, $\mathbf{R}'$ is not an optimal transmit covariance (i.e. one maximizing the secrecy rate $C(\mathbf{R})$) and the latter can only be found from $\max_{\mathbf{R} \in S_{\mathbf{R}}} C(\mathbf{R})$ and not from $\max_{\mathbf{R} \in S_{\mathbf{R}}} \min_{\mathbf{K} \in S_{\mathbf{K}}} f(\mathbf{R}, \mathbf{K})$ in this case, even though both problems have the same value. This does not happen without the IPC or if $\mathbf{W}_3$ is non-singular, i.e. both problems have the same optimal covariance, which is also the case in [18]. Note that $C(\mathbf{R}') = f(\mathbf{R}', \mathbf{K}')$ and $\mathbf{R}' = \mathbf{R}^*$ is unique and optimal when $P_T \leq P_I$ in this example, i.e. when the TPC is active, as expected from Proposition 16.

The next example demonstrates that while unbounded grows of C_s with P_T is possible without the IPC, this may be drastically changed when the IPC is introduced. Let us consider the following channel:

$$\mathbf{W}_1 = \begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.5 \end{pmatrix}, \quad \mathbf{W}_2 = \begin{pmatrix} 0.4 & -0.4 \\ -0.4 & 0.4 \end{pmatrix},$$
$$\mathbf{W}_{3a} = \begin{pmatrix} 0.1 & 0.1 \\ 0.1 & 0.1 \end{pmatrix}, \quad \mathbf{W}_{3b} = \begin{pmatrix} 0.1 & -0.1 \\ -0.1 & 0.1 \end{pmatrix} \quad (56)$$

for which the secrecy capacity is shown in Fig. 10. While the capacity growths unbounded with P_T without the IPC (since $r(\mathbf{W}_2) = 1$ while $r(\mathbf{W}_1) = 2$, so that (40) is satisfied), it saturates with the IPC matrix $\mathbf{W}_{3a}$ but not with $\mathbf{W}_{3b}$. Hence, the introduction of the IPC may have a dramatic impact on the high-SNR behaviour of the secrecy capacity and this also depends significantly on the specifics of the IPC matrix, including the phases of its off-diagonal terms. Note also that the low-SNR behaviour is not affected by the IPC while the high-SNR behaviour may or may not be affected, depending on $\mathbf{W}_3$.

X. Conclusion

The Gaussian MIMO wiretap channel was studied under interference constraints. Its operational secrecy capacity has been established in two equivalent forms, as a non-convex max problem and a convex-concave max-min problem. While no closed-form solution is known to any of these problems in the general case (even without interference constraints), rank-1 and high-rank solutions have been obtained under zero-interference constraints. Optimal signaling directions have been characterized in the general case, from which tight bounds to the rank of optimal covariance were obtained. Sufficient condition for the optimality of beamforming as well as necessary condition for full-rank optimal signaling were given. The results are extended to a multi-user scenario. Sufficient and necessary conditions for non-zero secrecy capacity and its unbounded growth under interference constraints were given. The interplay between total transmit and interference power constraints was shown to affect the optimal signaling in a significant way, especially in the high-SNR regime. Overall, these results provide insights into fundamental information-theoretic limits as well as optimal signaling strategies for secure communications under interference constraints.

Appendix

A. Proof of Theorem 1

The proof is via the sequence of Propositions below and is based on the method of [17], [18] by properly extending it to incorporate the interference constraint. First, lower and upper bounds to the secrecy capacity are established as follows.

Proposition 11: Let $p_{\mathbf{x}}$ be a probability distribution of input $\mathbf{x}$. The secrecy capacity C_s of the CR MIMO WTC can be bounded as follows:

$$\max_{p_{\mathbf{x}} \in \mathcal{P}} [I(\mathbf{x}; \mathbf{y}_1) - I(\mathbf{x}; \mathbf{y}_2)] \leq C_s \leq \max_{p_{\mathbf{x}} \in \mathcal{P}} I(\mathbf{x}; \mathbf{y}_1 | \mathbf{y}_2) \quad (57)$$

where $I(\mathbf{x}; \mathbf{y}_{1(2)})$ is the mutual information between $\mathbf{x}$ and $\mathbf{y}_{1(2)}$, and $I(\mathbf{x}; \mathbf{y}_1 | \mathbf{y}_2)$ is the conditional mutual information when $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ are jointly Gaussian and the covariance *of $[\boldsymbol{\xi}_1^+, \boldsymbol{\xi}_2^+]^+$ is $\mathbf{K}$ (as in* (6)*); $\mathcal{P}$ is the set of all input distributions $p_{\mathbf{x}}$ that satisfy the TPC and IPC:*

$$\mathcal{P} = \{p_{\mathbf{x}} : E\left\{|\mathbf{x}|^2\right\} \le P_T,\ E\left\{|\mathbf{H}_3\mathbf{x}|^2\right\} \le P_I\}. \tag{58}$$

Proof: The upper bound in (57) is obtained via a genie-aided channel in which the Rx observes $\mathbf{y}_2$ in addition to $\mathbf{y}_1$. Such channel has in general a larger or equal capacity to that of the original channel and it is degraded at the same time, making the analysis much simpler (since the original Wyner's construction of the converse applies). Furthermore, one can always choose the noises $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ to be jointly Gaussian and correlated with each other (since the secrecy capacity depends on the marginal distributions, not the joint one [15]) and select their cross-covariance as to minimize the upper bound as in [17]. The details follow below.

To establish the upper bound, consider a $(2^{nR}, n)$ code for the channel, where n is the blocklength and R is the rate, so that a message w is uniformly distributed over the set $\left\{1, 2, \ldots, 2^{nR}\right\}$. An encoder maps the message w to the transmitted vector sequence $\{\mathbf{x}(t)\}_{t=1}^n$, and a decoder maps the received sequence $\{\mathbf{y}(t)\}_{t=1}^n$ to a message estimate $\hat{w}$. Let

$$\begin{aligned}\mathbf{X}^n &= [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(n)],\\ \mathbf{Y}_i^n &= [\mathbf{y}_i(1), \mathbf{y}_i(2), \ldots, \mathbf{y}_i(n)], \quad i = 1, 2\end{aligned}$$

denote the transmitted and received sequence matrix from time 1 to n respectively. The reliability and secrecy criteria are as follows: for every $\epsilon > 0$ and n sufficiently large,

$$\Pr(w \ne \hat{w}) \le \epsilon, \quad n^{-1} I(w; \mathbf{Y}_2^n) \le \epsilon, \tag{59}$$

while the TPC and IPC are

$$\frac{1}{n}\sum_{i=1}^n E\left\{|\mathbf{x}(i)|^2\right\} \le P_T, \tag{60}$$

$$\frac{1}{n}\sum_{i=1}^n E\left\{|\mathbf{H}_3\mathbf{x}(i)|^2\right\} \le P_I. \tag{61}$$

Note that (59) implies, from Fano's inequality, that

$$n^{-1} I(w; \mathbf{Y}_1^n) \geqslant R - \epsilon_F \tag{62}$$

where $\epsilon_F \to 0$ as $\epsilon \to 0$. By combining (59) and (62), we obtain, for any $\epsilon' = \epsilon_F + \epsilon > 0$,

$$\begin{aligned}R - \epsilon' &\leqslant n^{-1}[I(w; \mathbf{Y}_1^n) - I(w; \mathbf{Y}_2^n)]\\ &\leqslant n^{-1}[h(\mathbf{Y}_1^n|\mathbf{Y}_2^n) - h(\mathbf{Y}_1^n|\mathbf{Y}_2^n, w, \mathbf{X}^n)] && (63)\\ &\leqslant n^{-1}\sum_{i=1}^n I(\mathbf{x}(i); \mathbf{y}_1(i)|\mathbf{y}_2(i)) && (64)\\ &\leqslant I(\bar{\mathbf{x}}_q; \bar{\mathbf{y}}_{1q}|\bar{\mathbf{y}}_{2q}) && (65)\\ &\le \max_{p_{\mathbf{x}} \in \mathcal{P}} I(\mathbf{x}; \mathbf{y}_1|\mathbf{y}_2) && (66)\end{aligned}$$

where (63)-(66) are obtained via the same steps as in [17, Appendix I]. In (65), $q = 1, 2, \ldots, n$ is a time-sharing random variable with uniform distribution: $p_q = 1/n$, and $\bar{\mathbf{x}}_q$ is the composite (time-shared) random input whose distribution is the average of those of $\mathbf{x}(1), .., \mathbf{x}(n)$:

$$p_{\bar{\mathbf{x}}_q} = n^{-1}\sum_{i=1}^n p_{\mathbf{x}(i)}. \tag{67}$$

To prove (66), note that since $\{\mathbf{x}(1), \ldots, \mathbf{x}(n)\}$ satisfy the power and interference constrained in (60) and (61), so is $\bar{\mathbf{x}}_q$, since

$$\begin{aligned}n^{-1}\sum_{i=1}^n E\left\{|\mathbf{x}(i)|^2\right\} &= \sum_{i=1}^n p(q = i) E\left\{|\mathbf{x}_q|^2 | q = i\right\}\\ &= E_{\mathbf{x}|q}\left\{|\mathbf{x}_q|^2 | q\right\}\\ &= E\left\{|\bar{\mathbf{x}}_q|^2\right\}\\ &\le P_T, && (68)\\ n^{-1}\sum_{i=1}^n E\left\{|\mathbf{H}_3\mathbf{x}(i)|^2\right\} &= \sum_{i=1}^n p(q = i) E\left\{|\mathbf{H}_3\mathbf{x}_q|^2 | q = i\right\}\\ &= E_{\mathbf{x}|q}\left\{|\mathbf{H}_3\mathbf{x}_q|^2 | q\right\}\\ &= E\left\{|\mathbf{H}_3\bar{\mathbf{x}}_q|^2\right\}\\ &\le P_I && (69)\end{aligned}$$

holds. Thus, the desired upper bound in (66) follows. Since this holds for any $\epsilon' > 0$, the upper bound in (57) follows.

To establish the lower bound in (57), note that it is an achievable rate, which follows from the Csiszar-Korner formula (see e.g. [15], [17]) by setting $\mathbf{u} = \mathbf{x}$ and using an input subject to the constraints in (60) and (61), where where $\mathbf{u}$ is the auxiliary random variable in the Csiszar-Korner formula. □

Note that this proposition does not require $\mathbf{x}$ to be Gaussian. The following proposition establishes the optimality of Gaussian inputs.

Proposition 12: For each $\mathbf{K} > \mathbf{0}$, the distribution of $\mathbf{x}$ maximizing $I(\mathbf{x}; \mathbf{y}_1|\mathbf{y}_2)$ in (57) *is Gaussian.*

Proof: $I(\mathbf{x}; \mathbf{y}_1|\mathbf{y}_2)$ can be expressed as

$$\begin{aligned}I(\mathbf{x}; \mathbf{y}_1|\mathbf{y}_2) &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\mathbf{y}_1|\mathbf{y}_2, \mathbf{x})\\ &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1|\mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2, \mathbf{x})\\ &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\boldsymbol{\xi}_1|\boldsymbol{\xi}_2, \mathbf{x})\\ &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\boldsymbol{\xi}_1|\boldsymbol{\xi}_2) && (70)\end{aligned}$$

Since the second term in (70) is independent of $\mathbf{x}$, it suffices to establish that $h(\mathbf{y}_1|\mathbf{y}_2)$ is maximized when $\mathbf{x}$ is Gaussian. While Gaussian distribution maximises the differential entropy under covariance constraint, it is not necessarily so for conditional entropy, since it is a difference of 2 differential entropies. To this end, we need the following Lemma of Thomas [29].

Lemma 3: Let $z_1, .., z_k$ be a set of arbitrary zero-mean random variables with covariance matrix $\mathbf{R}$. Let S be any subset of $\{1, 2, \ldots, k\}$ and $\bar{S}$ be its complement. Then

$$h(\mathbf{z}_S|\mathbf{z}_{\bar{S}}) \le h(\mathbf{z}_S^*|\mathbf{z}_{\bar{S}}^*) \tag{71}$$

where $(z_1^*, .., z_k^*) \sim N(0, \mathbf{R})$, i.e. Gaussian with the same mean and covariance.

Since $\mathbf{R}$ in this Lemma is arbitrary, the TPC and IPC can be accommodated by properly restricting the choice of $\mathbf{R}$. Applying this inequality to $h(\mathbf{y}_1|\mathbf{y}_2)$ and maximizing the upper bound over $\mathbf{R} \in S_{\mathbf{R}}$, one concludes that Gaussian input achieves the upper bound in (57), since, under such input, $\mathbf{y}_1, \mathbf{y}_2$ are also Gaussian. □

Since Gaussian input maximizes the upper bound, $I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)$ under such input can be expressed as

$$\begin{aligned} I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2) &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\mathbf{y}_1|\mathbf{x},\mathbf{y}_2) \\ &= h(\mathbf{y}_1|\mathbf{y}_2) - h(\mathbf{x},\mathbf{y}_1,\mathbf{y}_2) + h(\mathbf{x},\mathbf{y}_2) \\ &= h(\mathbf{y}_1,\mathbf{y}_2) - h(\mathbf{y}_2) - h(\mathbf{y}_1,\mathbf{y}_2|\mathbf{x}) + h(\mathbf{y}_2|\mathbf{x}) \\ &= \ln|\mathbf{I}+\mathbf{K}^{-1}\mathbf{H}\mathbf{R}\mathbf{H}^+| - \ln|\mathbf{I}+\mathbf{H}_2\mathbf{R}\mathbf{H}_2^+| \\ &= f(\mathbf{R},\mathbf{K}) \end{aligned} \tag{72}$$

Gaussian input can also be used for the lower bound and $\max_{p_\mathbf{x}}$ can be replaced by $\max_\mathbf{R}$ on the both sides of (57) (still preserving the inequalities), giving

$$\max_\mathbf{R} C(\mathbf{R}) \le C_s \le \max_\mathbf{R} f(\mathbf{R},\mathbf{K}) \tag{73}$$

which holds for any $\mathbf{K}$ and hence

$$\max_\mathbf{R} C(\mathbf{R}) \le C_s \le \min_\mathbf{K}\max_\mathbf{R} f(\mathbf{R},\mathbf{K}) \tag{74}$$

We further establish the existence of a saddle-point in the minimax problem above, which is essential to establish Propositions 14 and 15 below, on which Proposition 16 depends.

Proposition 13: The max-min problem in (8) *has a saddle point solution* $(\mathbf{R}',\mathbf{K}')$ *as in* (9) *and* (10)*.*

Proof: It was shown in [18] that $f(\mathbf{R},\mathbf{K})$ is concave in $\mathbf{R}$ for any fixed $\mathbf{K}$ and convex in $\mathbf{K}$ for any fixed $\mathbf{R}$. Since the feasible set $S_\mathbf{R}$ in (3) and $S_\mathbf{K}$ are convex (as an intersection of convex sets representing each constraint individually), von Neumann mini-max theorem applies (see e.g. [30]), from which (9) and hence (10) follow. □

Armed with the saddle-point solution, we further establish that it also solves the following entropy maximization problem.

Proposition 14: Let $h(\mathbf{y})$ *be the differential entropy of* $\mathbf{y}$ *and let* $\mathbf{Z}_{12}'$ *be the optimal minimum mean square error (MMSE) weight matrix to estimate* $\mathbf{y}_1$ *from* $\mathbf{y}_2$ *at saddle-point* $(\mathbf{R}',\mathbf{K}')$*:*

$$\mathbf{Z}_{12}' = (\mathbf{N}' + \mathbf{H}_1\mathbf{R}'\mathbf{H}_2^+)(\mathbf{I}+\mathbf{H}_2\mathbf{R}'\mathbf{H}_2^+)^{-1} \tag{75}$$

Then,

$$\arg\max_{\mathbf{R}\in S_R} h(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2) = \arg\max_{\mathbf{R}\in S_R} f(\mathbf{R},\mathbf{K}') \tag{76}$$

where $h(\cdot)$ *is evaluated under* $\mathbf{K}=\mathbf{K}'$*.*

Proof: First, we note that one cannot use the respective result from [18] directly since our feasible set $S_\mathbf{R}$ is different from that in [18] (in particular, it is *not* isotropic) so that the respective KKT conditions are different and this affects a number of key steps in [18]. Nevertheless, we demonstrate below that this important property does hold under the extra IPC.

After some manipulations, $h(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2)$ can be expressed as follows:

$$\begin{aligned} &h(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2) \\ &\quad= \ln|E\left\{(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2)(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2)^+\right\}| + n_1\ln(2\pi e) \\ &\quad= \ln|\mathbf{I}+\mathbf{B}_1+\mathbf{B}_2\mathbf{R}\mathbf{B}_2^+| + n_1\ln(2\pi e) \end{aligned} \tag{77}$$

where

$$\mathbf{B}_1 = \mathbf{Z}_{12}'\mathbf{Z}_{12}'^+ - \mathbf{Z}_{12}'\mathbf{N}'^+ - \mathbf{N}'\mathbf{Z}_{12}'^+,\ \mathbf{B}_2 = \mathbf{H}_1 - \mathbf{Z}_{12}'\mathbf{H}_2$$

and where the last term in (77) can be neglected in optimization. Note that $h(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2)$ is concave in $\mathbf{R}$, since log-det is a concave function and $\mathbf{I}+\mathbf{B}_1+\mathbf{B}_2\mathbf{R}\mathbf{B}_2^+$ is affine in $\mathbf{R}$. Since $h(\mathbf{y}_1 - \mathbf{Z}_{12}'\mathbf{y}_2)$ is concave in $\mathbf{R}$, the feasible set $S_\mathbf{R}$ is convex and Slater's condition holds, the KKT conditions are both necessary and sufficient for optimality of the LHS of (76), which take the following form:

$$\mathbf{B}_2^+[\mathbf{I}+\mathbf{B}_1+\mathbf{B}_2\mathbf{R}\mathbf{B}_2^+]^{-1}\mathbf{B}_2 + \mathbf{M}_1 - \lambda\mathbf{I} - \lambda_3\mathbf{W}_3 = 0 \tag{78}$$

$$\lambda(tr\mathbf{R} - P_T) = 0,\quad \lambda_3(tr\mathbf{W}_3\mathbf{R} - P_I) = 0, \tag{79}$$

$$\mathbf{M}_1\mathbf{R} = 0,\quad \mathbf{M}_1 \ge \mathbf{0},\ \lambda \ge 0,\ \lambda_3 \ge 0 \tag{80}$$

where $\mathbf{M}_1$ is a Lagrange multiplier responsible for the positive semi-definite constraints $\mathbf{R} \ge \mathbf{0}$, λ and λ_3 are Lagrange multiplier responsible for the total power $tr\mathbf{R} \le P_T$ and interference $tr\mathbf{W}_3\mathbf{R} \le P_I$ constraints.

Likewise, since $f(\mathbf{R},\mathbf{K}')$ is concave in $\mathbf{R}$, the KKT conditions are both necessary and sufficient for the optimality of the RHS of (76). After some manipulations (using matrix inversion Lemma etc.), they take the following form:

$$\mathbf{B}_2^+\mathbf{A}_{12}^{-1}\mathbf{B}_2 + \mathbf{M}_2 - \mu\mathbf{I} - \mu_3\mathbf{W}_3 = 0, \tag{81}$$

$$\mu(tr\mathbf{R} - P_T) = 0,\quad \mu_3(tr\mathbf{W}_3\mathbf{R} - P_I) = 0, \tag{82}$$

$$\mathbf{M}_2\mathbf{R} = 0,\quad \mathbf{M}_2 \ge \mathbf{0},\ \mu \ge 0,\ \mu_3 \ge 0 \tag{83}$$

where $\mathbf{M}_2$, μ, μ_3 are Lagrange multipliers and

$$\mathbf{A}_{12} = \mathbf{I} + \mathbf{H}_1\mathbf{R}\mathbf{H}_1^+ - \mathbf{Z}_{12}(\mathbf{N}'^+ + \mathbf{H}_2\mathbf{R}\mathbf{H}_1^+), \tag{84}$$

$$\mathbf{Z}_{12} = (\mathbf{N}' + \mathbf{H}_1\mathbf{R}\mathbf{H}_2^+)(\mathbf{I}+\mathbf{H}_2\mathbf{R}\mathbf{H}_2^+)^{-1}. \tag{85}$$

Note that $\mathbf{A}_{12}$ can be further expressed as

$$\begin{aligned} \mathbf{A}_{12} &= \mathbf{I} + \mathbf{H}_1\mathbf{R}\mathbf{H}_1^+ - \mathbf{Z}_{12}(\mathbf{N}'^+ + \mathbf{H}_2\mathbf{R}\mathbf{H}_1^+) \\ &\quad - (\mathbf{N}' + \mathbf{H}_1\mathbf{R}\mathbf{H}_2^+)\mathbf{Z}_{12}^+ + \mathbf{Z}_{12}(\mathbf{I}+\mathbf{H}_2\mathbf{R}\mathbf{H}_2^+)\mathbf{Z}_{12}^+ \\ &= \mathbf{I} + \mathbf{Z}_{12}\mathbf{Z}_{12}^+ - \mathbf{Z}_{12}\mathbf{N}'^+ - \mathbf{N}'\mathbf{Z}_{12}^+ \\ &\quad + (\mathbf{H}_1 - \mathbf{Z}_{12}\mathbf{H}_2)\mathbf{R}(\mathbf{H}_1 - \mathbf{Z}_{12}\mathbf{H}_2)^+ \\ &= \mathbf{I} + \mathbf{B}_1 + \mathbf{B}_2\mathbf{R}\mathbf{B}_2^+. \end{aligned} \tag{86}$$

so that the condition in (81) takes the form:

$$\mathbf{B}_2^+(\mathbf{I}+\mathbf{B}_1+\mathbf{B}_2\mathbf{R}\mathbf{B}_2^+)^{-1}\mathbf{B}_2 + \mathbf{M}_2 - \mu\mathbf{I} - \mu_3\mathbf{W}_3 = 0 \tag{87}$$

By comparing (78)-(80) to (82), (83) and (87), it is clear that any solution of the 1st set of KKT conditions also solves the 2nd one and hence optimal $\mathbf{R}$ are the same, as desired. □

Next, we consider the case when either the TPC is active or/and $\mathbf{W}_3$ is non-singular, so that $\mathbf{W}_\mu = \mu\mathbf{I} + \mu_3\mathbf{W}_3 > 0$, and deal with the singular case later on. Using Proposition 14, we establish the following property of the saddle-point under the stated conditions, which is needed to prove Proposition 16.

Proposition 15: If either the TPC is active or/and $\mathbf{W}_3$ *is non-singular, and* $\mathbf{H}_1 \ne \mathbf{Z}_{12}'\mathbf{H}_2$*, then the saddle point* $(\mathbf{R}',\mathbf{K}')$ *satisfies*

$$(\mathbf{N}'^+\mathbf{H}_1 - \mathbf{H}_2)\mathbf{S}' = 0 \tag{88}$$

for a full column-rank matrix $\mathbf{S}'$ *such that* $\mathbf{R}' = \mathbf{S}'\mathbf{S}'^+$ *(the rank factorization; the columns of* $\mathbf{S}'$ *are the scaled eigenvectors of* $\mathbf{R}'$ *corresponding to strictly-positive eigenvalues,* $r(\mathbf{R}') = r(\mathbf{S}')$*).*

Proof: Following 2nd inequality in (10) and the steps of the proof in [18, Lemma 3], one obtains

$$\mathbf{B}_2\mathbf{S}'\mathbf{S}'^{+}(\mathbf{N}'^{+}\mathbf{H}_1-\mathbf{H}_2)^{+}=\mathbf{0} \tag{89}$$

Using Proposition 14,

$$\begin{aligned}\mathbf{R}' &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}} f(\mathbf{R},\mathbf{K}')\\ &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}} h(\mathbf{y}_1-\mathbf{Z}'_{12}\mathbf{y}_2)\\ &= \arg\max_{\mathbf{R}\in S_{\mathbf{R}}} \ln|\mathbf{I}+\mathbf{H}_e\mathbf{R}\mathbf{H}_e^{+}| \end{aligned}\tag{90}$$

where $\mathbf{H}_e=(\mathbf{I}+\mathbf{B}_1)^{-1/2}\mathbf{B}_2$. At this point, we note that [18, Lemma 4] cannot be used to complete the proof since its proof makes use of the isotropy of the feasible set in [18], which is not the case here due to the IPC. Hence, a new line of attack is necessary. We proceed as follows.

Note that the equality in (90) is the following optimization problem (P1):

$$\begin{aligned}(\text{P1}):\ &\max_{\mathbf{R}}\ \ln|\mathbf{I}+\mathbf{W}_e\mathbf{R}|,\\ &\text{s.t. } \mathbf{R}\geq\mathbf{0}, tr\mathbf{R}\leq P_T,\quad tr(\mathbf{W}_3\mathbf{R})\leq P_I.\end{aligned}\tag{91}$$

where $\mathbf{W}_e=\mathbf{H}_e^{+}\mathbf{H}_e$, for which the KKT conditions are

$$(\mathbf{I}+\mathbf{W}_e\mathbf{R})\mathbf{W}_\mu=\mathbf{W}_e+\mathbf{M}, \tag{92}$$

$$\begin{aligned}&\mathbf{M}\mathbf{R}=\mathbf{0},\quad \mu(tr\mathbf{R}-P_T)=0,\\ &\mu_3(tr(\mathbf{W}_3\mathbf{R})-P_I)=0,\end{aligned}\tag{93}$$

$$\begin{aligned}&tr\mathbf{R}\leq P_T,\quad tr(\mathbf{W}_3\mathbf{R})\leq P_I,\ \mathbf{R}\geq\mathbf{0},\ \mathbf{M}\geq\mathbf{0},\\ &\mu\geq 0,\quad \mu_3\geq 0\end{aligned}\tag{94}$$

where $\mu,\ \mu_3,\ \mathbf{M}\geq 0$ are Lagrangian multipliers responsible for $tr\mathbf{R}\leq P_T,\ tr(\mathbf{W}_3\mathbf{R})\leq P_I,\ \mathbf{R}\geq\mathbf{0}$ respectively and where $\mathbf{W}_\mu=\mu\mathbf{I}+\mu_3\mathbf{W}_3$. Multiplying both sides of (92) by $\mathbf{W}_\mu^{-\frac{1}{2}}$ from the left and the right, one obtains

$$\tilde{\lambda}(\mathbf{I}+\tilde{\mathbf{W}}_e\tilde{\mathbf{R}})=\tilde{\mathbf{W}}_e+\tilde{\mathbf{M}} \tag{95}$$

where $\tilde{\lambda}=1$, $\tilde{\mathbf{W}}_e=\mathbf{W}_\mu^{-\frac{1}{2}}\mathbf{W}_e\mathbf{W}_\mu^{-\frac{1}{2}}$, $\tilde{\mathbf{R}}=\mathbf{W}_\mu^{\frac{1}{2}}\mathbf{R}\mathbf{W}_\mu^{\frac{1}{2}}$, $\tilde{\mathbf{M}}=\mathbf{W}_\mu^{-\frac{1}{2}}\mathbf{M}\mathbf{W}_\mu^{-\frac{1}{2}}$, and hence

$$\tilde{\mathbf{M}}\tilde{\mathbf{R}}=\mathbf{0},\quad \tilde{\mathbf{R}}\geq\mathbf{0},\ \tilde{\mathbf{M}}\geq\mathbf{0} \tag{96}$$

Define $P'_T=tr(\mathbf{W}_\mu^{\frac{1}{2}}\mathbf{R}'\mathbf{W}_\mu^{\frac{1}{2}})$. Then, (95)-(96), in combination with

$$\tilde{\lambda}(tr\tilde{\mathbf{R}}-P'_T)=0 \tag{97}$$

are just the KKT conditions of the following problem (P2):

$$\begin{aligned}(\text{P2}):\ &\max_{\tilde{\mathbf{R}}}\ \ln|\mathbf{I}+\tilde{\mathbf{W}}_e\tilde{\mathbf{R}}|,\\ &\text{s.t. } \tilde{\mathbf{R}}\geq\mathbf{0},\ tr\tilde{\mathbf{R}}\leq P'_T.\end{aligned}\tag{98}$$

so that any solution of (P1) is also a solution of (P2) (since both (P1) and (P2) are convex problems and Slater's condition holds, their KKT conditions are sufficient and necessary for optimality [30]). However, the feasible set of (P2) is isotropic (unlike that of (P1)) and hence [18, Lemma 4] applies to (P2) so that $\tilde{\mathbf{H}}_e\tilde{\mathbf{S}}'$ is full column-rank matrix[3], where $\tilde{\mathbf{H}}_e=\mathbf{H}_e\mathbf{W}_\mu^{-\frac{1}{2}}$ and $\tilde{\mathbf{R}}'=\tilde{\mathbf{S}}'\tilde{\mathbf{S}}'^{+}$ is a solution of (P2) with $\tilde{\mathbf{S}}'$ being full-column rank. Further note that

$$\begin{aligned}\tilde{\mathbf{H}}_e\tilde{\mathbf{S}}' &= \mathbf{H}_e\mathbf{W}_\mu^{-\frac{1}{2}}\mathbf{W}_\mu^{\frac{1}{2}}\mathbf{S}'\\ &= \mathbf{H}_e\mathbf{S}'\\ &= (\mathbf{I}+\mathbf{B}_1)^{-1/2}\mathbf{B}_2\mathbf{S}'\end{aligned}\tag{99}$$

and hence $\mathbf{B}_2\mathbf{S}'$ is of full column-rank so that

$$\mathbf{S}'^{+}(\mathbf{N}'^{+}\mathbf{H}_1-\mathbf{H}_2)^{+}=\mathbf{0} \tag{100}$$

according to (89), from which (88) follows. □

The final step is to show that the upper bound and lower bounds in (74) coincide.

Proposition 16: If either the TPC is active or/and $\mathbf{W}_3$ *is non-singular, the saddle point solution* $(\mathbf{R}',\mathbf{K}')$ *satisfies*

$$f(\mathbf{R}',\mathbf{K}')=C(\mathbf{R}') \tag{101}$$

Proof: We consider 1st the case of $\mathbf{H}_1\neq\mathbf{Z}'_{12}\mathbf{H}_2$. To this end, take Gaussian $\mathbf{x}$ and use the chain rule to obtain

$$f(\mathbf{R}',\mathbf{K}')=I(\mathbf{x};\mathbf{y}_1|\mathbf{y}_2)=C(\mathbf{R}')+I(\mathbf{x};\mathbf{y}_2|\mathbf{y}_1) \tag{102}$$

Using (70), one can express $I(\mathbf{x};\mathbf{y}_2|\mathbf{y}_1)$ as

$$\begin{aligned}I(\mathbf{x},\mathbf{y}_2|\mathbf{y}_1)=h(\mathbf{y}_2-\mathbf{Z}_{21}\mathbf{y}_1)-\ln|\mathbf{I}-\mathbf{N}'^{+}\mathbf{N}'|\\ -n_2\ln(2\pi e)\end{aligned}\tag{103}$$

where

$$\mathbf{Z}_{21}=(\mathbf{N}'^{+}+\mathbf{H}_2\mathbf{R}'\mathbf{H}_1^{+})(\mathbf{I}+\mathbf{H}_1\mathbf{R}'\mathbf{H}_1^{+})^{-1} \tag{104}$$

denotes the MMSE matrix of estimating $\mathbf{y}_2$ from $\mathbf{y}_1$ and $\mathbf{N}^{+}$ also represents the MMSE matrix of estimating $\boldsymbol{\xi}_2$ from $\boldsymbol{\xi}_1$.

At a saddle point $(\mathbf{R}',\mathbf{K}')$, one obtains:

$$\begin{aligned}&h(\mathbf{y}_2-\mathbf{Z}_{21}\mathbf{y}_1)-n_2\ln(2\pi e)\\ &\quad=\ln|\mathbf{I}+\mathbf{H}_2\mathbf{R}'\mathbf{H}_2^{+}-\mathbf{Z}_{21}(\mathbf{N}'+\mathbf{H}_1\mathbf{R}'\mathbf{H}_2^{+})|\\ &\quad=\ln|\mathbf{I}+\mathbf{H}_2\mathbf{R}'\mathbf{H}_2^{+}-\mathbf{N}'^{+}(\mathbf{I}+\mathbf{H}_1\mathbf{S}'\mathbf{S}'^{+}\mathbf{H}_1^{+})\mathbf{N}'| \qquad (105)\\ &\quad=\ln|\mathbf{I}+\mathbf{N}'^{+}\mathbf{H}_1\mathbf{S}'\mathbf{S}'^{+}\mathbf{H}_1^{+}\mathbf{N}'\\ &\qquad-\mathbf{N}'^{+}(\mathbf{I}+\mathbf{H}_1\mathbf{S}'\mathbf{S}'^{+}\mathbf{H}_1^{+})\mathbf{N}'|\\ &\quad=\ln|\mathbf{I}-\mathbf{N}'^{+}\mathbf{N}'| \qquad (106)\end{aligned}$$

where (105) and (106) are obtained via $\mathbf{N}'^{+}\mathbf{H}_1\mathbf{S}'=\mathbf{H}_2\mathbf{S}'$ from Proposition 15. Therefore, using (102) and (103),

$$I(\mathbf{x},\mathbf{y}_2|\mathbf{y}_1)=f(\mathbf{R}',\mathbf{K}')-C(\mathbf{R}')=0. \tag{107}$$

Thus, $f(\mathbf{R}',\mathbf{K}')=C(\mathbf{R}')$ as desired.

Finally, we show that $f(\mathbf{R}',\mathbf{K}')=0$ if $\mathbf{H}_1=\mathbf{Z}'_{12}\mathbf{H}_2$. To this end, note that

$$\mathbf{y}_1-\mathbf{Z}'_{12}\mathbf{y}_2=\boldsymbol{\xi}_1-\mathbf{Z}'_{12}\boldsymbol{\xi}_2. \tag{108}$$

so that

$$f(\mathbf{R}',\mathbf{K}')=h(\boldsymbol{\xi}_1-\mathbf{Z}'_{12}\boldsymbol{\xi}_2)-h(\boldsymbol{\xi}_1-\mathbf{N}'\boldsymbol{\xi}_2). \tag{109}$$

Substituting $\mathbf{H}_1=\mathbf{Z}'_{12}\mathbf{H}_2$ into (75) and after some manipulations, one obtains

$$\mathbf{Z}'_{12}(\mathbf{I}+\mathbf{H}_2\mathbf{R}'\mathbf{H}_2^{+})=(\mathbf{N}'+\mathbf{Z}'_{12}\mathbf{H}_2\mathbf{R}'\mathbf{H}_2^{+}). \tag{110}$$

so that $\mathbf{Z}'_{12}=\mathbf{N}'$ and hence $f(\mathbf{R}',\mathbf{K}')=0$. But $C(\mathbf{R}')\leq f(\mathbf{R}',\mathbf{K}')$, so that $C(\mathbf{R}')=0$, as desired. □

[3] This can also be seen using the standard WF solution for (P2).

Note that, from (101), $\mathbf{R}'$ is also a maximizer of $C(\mathbf{R})$ when $\mathbf{W}_\mu$ is not singular, $\mathbf{R}^* = \mathbf{R}'$, which is not necessarily true for a singular $\mathbf{W}_\mu$.

Combining (74) and (101), Theorem 1 follows. While we considered here the case of non-singular $\mathbf{K}$ only, the singular case can be established in a similar way with somewhat more lengthy arguments (using pseudo-inverse instead of the inverse and related projection on the active sub-space only).

Since Proposition 15 was established for non-singular $\mathbf{W}_\mu$ only, i.e. when either the TPC is active or/and $\mathbf{W}_3$ is non-singular, the above result is limited to this case only (in fact, it can be shown, by examples, that it does not hold in the singular case). Let us now consider the case of singular $\mathbf{W}_\mu$. Note that a singular $\mathbf{W}_\mu$ implies $\mu = 0$ (inactive TPC) and singular $\mathbf{W}_3$. We deal with it using the standard continuity argument: let $\mathbf{W}_3 \rightarrow \mathbf{W}_{3\epsilon} = \mathbf{W}_3 + \epsilon\mathbf{I}$ for some $\epsilon > 0$, so that

$$\mathbf{W}_\mu = \mu_3(\mathbf{W}_3 + \epsilon\mathbf{I}) > 0 \tag{111}$$

and observe that Proposition 15 and hence Proposition 16 now apply for any $\epsilon > 0$. Next, we need the following continuity properties, which are straightforward to prove (using the continuity of the objective functions and compactness of the feasible sets).

Lemma 4: Consider the following "regularized" problems

$$C_\epsilon = \max_{\mathbf{R}\in S_{\mathbf{R}\epsilon}} C(\mathbf{R}), \tag{112}$$

$$f_\epsilon = \max_{\mathbf{R}\in S_{\mathbf{R}\epsilon}} \min_{\mathbf{K}} f(\mathbf{R},\mathbf{K}) \tag{113}$$

where $S_{\mathbf{R}\epsilon} = \{\mathbf{R} \geq \mathbf{0} : tr\mathbf{R} \leq P_T,\ tr(\mathbf{W}_{3\epsilon}\mathbf{R}) \leq P_I\}$. *Then,*

$$\lim_{\epsilon\rightarrow 0} C_\epsilon = \max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R}) \tag{114}$$

$$\lim_{\epsilon\rightarrow 0} f_\epsilon = \max_{\mathbf{R}\in S_{\mathbf{R}}} \min_{\mathbf{K}} f(\mathbf{R},\mathbf{K}) = f(\mathbf{R}',\mathbf{K}') \tag{115}$$

Now observe that

$$C_\epsilon = f_\epsilon \quad \forall\epsilon > 0 \tag{116}$$

due to Proposition 16, and take the limit $\epsilon \rightarrow 0$ to establish

$$\max_{\mathbf{R}\in S_{\mathbf{R}}} C(\mathbf{R}) = \max_{\mathbf{R}\in S_{\mathbf{R}}} \min_{\mathbf{K}} f(\mathbf{R},\mathbf{K}) = f(\mathbf{R}',\mathbf{K}') \tag{117}$$

in the singular case. We remark that $C(\mathbf{R}')$ does not have to be equal to $f(\mathbf{R}',\mathbf{K}')$ in the singular case, i.e. $\mathbf{R}'$ is not necessarily a maximizer of $C(\mathbf{R})$ if $\mathbf{W}_\mu$ is singular.

B. Proof of Proposition 1

The KKT conditions for 1st problem in (8) are

$$(\mathbf{I}+\mathbf{W}_1\mathbf{R})^{-1}\mathbf{W}_1 - \mathbf{W}_2(\mathbf{I}+\mathbf{R}\mathbf{W}_2)^{-1} + \mathbf{M} = \mathbf{W}_\mu \tag{118}$$

$$\mu(tr\mathbf{R} - P_T) = 0, \quad \mu_3(tr\mathbf{W}_3\mathbf{R} - P_I) = 0,$$

$$\mathbf{R}\mathbf{M} = 0 \tag{119}$$

$$\mu,\ \mu_3 \geq 0, \quad \mathbf{R},\mathbf{M} \geq 0 \tag{120}$$

$$tr\mathbf{R} \leq P_T, \quad tr\mathbf{W}_3\mathbf{R} \leq P_I \tag{121}$$

where $\mu,\ \mu_3 \geq 0$ are Lagrange multipliers responsible for the TPC and IPC respectively; $\mathbf{M} \geq 0$ is a (matrix) Lagrange multiplier responsible for the positive semi-definite constraint $\mathbf{R} \geq 0$; $\mathbf{W}_\mu = \mu\mathbf{I} + \mu_3\mathbf{W}_3$; (118) is the stationarity condition, (119) are the complementary slackness conditions, (120) and (121) are primal and dual feasibility conditions. While these conditions are not sufficient for optimality in the general (non-degraded) case, they are necessary, since the (affine) constraints $tr\mathbf{R} \leq P_T$, $tr\mathbf{W}_3\mathbf{R} \leq P_I$, $\mathbf{R} \geq 0$ clearly satisfy the Slater condition and since the maximum is achievable (since the constraint set is compact and the objective function is continuous) [31]. (118) can be expressed as

$$\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M} = (\mathbf{I}+\mathbf{W}_1\mathbf{R})\mathbf{W}_\mu(\mathbf{I}+\mathbf{R}\mathbf{W}_2) \tag{122}$$

Using $\mathbf{R} = \mathbf{Q}\mathbf{Q}^+$, where $\mathbf{Q} = \mathbf{U}_+\mathbf{\Lambda}_+^{1/2}$ and $\mathbf{\Lambda}_+$ is a diagonal matrix of strictly positive eigenvalues of $\mathbf{R}$, and multiplying both sides of (122) by $\mathbf{Q}^+$ from the left and by $\mathbf{Q}$ from the right, one obtains

$$\mathbf{Q}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{Q} = (\mathbf{I}+\mathbf{W}_1')\mathbf{W}_\mu'(\mathbf{I}+\mathbf{W}_2') \tag{123}$$

where $\mathbf{W}_k' = \mathbf{Q}^+\mathbf{W}_k\mathbf{Q}$. Further note that $\mathbf{I} + \mathbf{W}_k' > 0$, $\mathbf{W}_\mu' \geq 0$. We further need the following technical Lemma (see [24] for a proof).

Lemma 5: Let $\mathbf{A},\mathbf{B},\mathbf{C} \geq 0$ *be positive semi-definite matrices and let* $\mathbf{ABC}$ *be Hermitian. Then* $\mathbf{ABC} \geq 0$ *and the inequality is strict when all matrices are full-rank.*

Applying this Lemma to (123), one concludes

$$\mathbf{Q}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{Q} \geq 0 \tag{124}$$

from which (15) follows (since $\mathbf{\Lambda}_+ > 0$). When either the TPC is active or/and $\mathbf{W}_3 > 0$, then $\mathbf{W}_\mu > 0$ and the strict inequality follows by taking the determinant of both sides of (123).

C. Proof of Proposition 2

We need the following technical Lemma, which is a direct consequence of Corollary 4.5.11 in [32].

Lemma 6: Let $\mathbf{A}$ *be Hermitian and* $r_+(\mathbf{A})$ *be its number of positive eigenvalues. Then* $r_+(\mathbf{S}^+\mathbf{A}\mathbf{S}) \leq r_+(\mathbf{A})$, *where* $\mathbf{S}$ *is any matrix of appropriate size.*

To prove (18), note that when either TPC is active or/and $\mathbf{W}_3 > 0$, then $\mathbf{W}_\mu > 0$ and hence, form (122), $\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M}$ is full-rank, so that

$$\begin{aligned} r(\mathbf{R}^*) &= r(\mathbf{Q}(\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M})\mathbf{Q}^+) \\ &= r_+(\mathbf{Q}(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{Q}^+) \\ &\leq r_+(\mathbf{W}_1 - \mathbf{W}_2) \end{aligned} \tag{125}$$

as required, where $\mathbf{R}^* = \mathbf{Q}\mathbf{Q}^+$. 2nd inequality in (18) follows from $r_+(\mathbf{W}) + r_-(\mathbf{W}) \leq m$.

To prove (17), note that, from (122),

$$\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M} \geq 0 \tag{126}$$

in general. Expanding

$$\mathbf{W}_1 - \mathbf{W}_2 = \mathbf{W}_+ - \mathbf{W}_-,\ \mathbf{W}_- = \mathbf{U}_-\mathbf{\Lambda}_-\mathbf{U}_-^+ \tag{127}$$

where $\mathbf{W}_{+(-)}$ collect positive(negative) eigenmodes of $\mathbf{W}_1 - \mathbf{W}_2$, the columns of semi-unitary matrix $\mathbf{U}_-$ are the active eigenvectors of $\mathbf{W}_-$ and $\mathbf{\Lambda}_- > 0$ is a diagonal matrix of its eigenvalues, and multiplying (126) by $\mathbf{U}_-^+$ and $\mathbf{U}_-$, one obtains

$$\mathbf{U}_-^+(\mathbf{W}_1 - \mathbf{W}_2 + \mathbf{M})\mathbf{U}_- = -\mathbf{\Lambda}_- + \mathbf{U}_-^+\mathbf{M}\mathbf{U}_- \geq 0 \quad (128)$$

from which it follows that

$$r(\mathbf{M}) \geq r(\mathbf{U}_-^+\mathbf{M}\mathbf{U}_-) \geq r(\mathbf{W}_-) = r_-(\mathbf{W}_1 - \mathbf{W}_2) \quad (129)$$

Since $\mathbf{R}^*\mathbf{M} = 0$, it follows that $\mathcal{R}(\mathbf{R}^*) \in \mathcal{N}(\mathbf{M})$ and hence

$$\begin{aligned} r(\mathbf{R}^*) &\leq \dim \mathcal{N}(\mathbf{M}) = m - r(\mathbf{M}) \\ &\leq m - r_-(\mathbf{W}_1 - \mathbf{W}_2) \end{aligned} \quad (130)$$

as required, where $\dim \mathcal{N}(\mathbf{M})$ is the dimensionality of $\mathcal{N}(\mathbf{M})$.

D. Proof of Proposition 3

Part 1: to prove sufficiency, observe that $r_+(\mathbf{W}_1 - \mathbf{W}_2) \geq 1$ implies that there exists unitary vector $\mathbf{u} : \ \mathbf{u}^+\mathbf{W}_1\mathbf{u} > \mathbf{u}^+\mathbf{W}_2\mathbf{u}$, so that one can set $\mathbf{R} = p\mathbf{u}\mathbf{u}^+$, where $p = \min[P_T, P_I/\mathbf{u}^+\mathbf{W}_3\mathbf{u}] > 0$, which is feasible, and

$$C_s \geq C(\mathbf{R}) = \ln \frac{1 + p\mathbf{u}^+\mathbf{W}_1\mathbf{u}}{1 + p\mathbf{u}^+\mathbf{W}_2\mathbf{u}} > 0 \quad (131)$$

To prove necessity, assume $r_+(\mathbf{W}_1 - \mathbf{W}_2) = 0$, which is equivalent to $\mathbf{W}_1 \leq \mathbf{W}_2$ and hence

$$C(\mathbf{R}) = \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} \leq 0 \quad (132)$$

for any $\mathbf{R} \geq 0$ so that $C_s = \max_{\mathbf{R}} C(\mathbf{R}) = 0$.

Part 2: sufficiency is proved similarly to that of Part 1. $r_+(\mathbf{U}_0^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_0) \geq 1$ implies

$$\exists \mathbf{u} : \ \mathbf{u}^+\mathbf{W}_1\mathbf{u} > \mathbf{u}^+\mathbf{W}_2\mathbf{u}, \quad \mathbf{u} \in \mathcal{N}(\mathbf{W}_3)$$

so that one can set $\mathbf{R} = P_T\mathbf{u}\mathbf{u}^+$, which is feasible (note that $tr\mathbf{W}_3\mathbf{R} = 0$, $tr\mathbf{R} = P_T$) and for which $C(\mathbf{R}) > 0$. To prove necessity, assume $r_+(\mathbf{U}_0^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_0) = 0$, which implies that $\mathbf{U}_0^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_0 \leq 0$. On the other hand, $tr\mathbf{W}_3\mathbf{R} = 0$ implies $\mathbf{W}_3\mathbf{R} = 0$, which implies $\mathcal{R}(\mathbf{R}) \in \mathcal{N}(\mathbf{W}_3)$ and hence $\mathbf{R}$ can be presented as $\mathbf{R} = \mathbf{U}_0\tilde{\mathbf{R}}\mathbf{U}_0^+$ for some $\tilde{\mathbf{R}} \geq 0$, and hence

$$C(\mathbf{R}) = \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} = \ln \frac{|\mathbf{I} + \tilde{\mathbf{W}}_1\tilde{\mathbf{R}}|}{|\mathbf{I} + \tilde{\mathbf{W}}_2\tilde{\mathbf{R}}|} \leq 0 \quad (133)$$

for any $\tilde{\mathbf{R}} \geq 0$ and hence for any feasible $\mathbf{R}$, where $\tilde{\mathbf{W}}_k = \mathbf{U}_0^+\mathbf{W}_k\mathbf{U}_0$, so that $C_s = \max_{\mathbf{R}} C(\mathbf{R}) = 0$.

E. Proof of Lemma 1

Since $\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+ \geq 0$, $\lambda_i(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) \geq 0$ so that

$$tr(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) = \sum_i \lambda_i(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) = 0 \quad (134)$$

imply $\lambda_i(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) = 0$ and hence $\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+ = 0$.

To show $\mathbf{H}_3\mathbf{R} = \mathbf{0}$, observe that

$$\sigma_i^2(\mathbf{H}_3\mathbf{R}^{1/2}) = \lambda_i(\mathbf{H}_3\mathbf{R}\mathbf{H}_3^+) = 0 \quad (135)$$

where σ_i denotes singular values, so that $\mathbf{H}_3\mathbf{R}^{1/2} = 0$ and hence $\mathbf{H}_3\mathbf{R} = 0$ and $\mathbf{W}_3\mathbf{R} = 0$ follow. Using similar approach, it can be shown that $\mathbf{W}_3\mathbf{R} = 0$ implies $\mathbf{H}_3\mathbf{R} = 0$, so that these conditions are equivalent.

F. Proof of Proposition 4

Let $\mathbf{R}^*$ and $\tilde{\mathbf{R}}^*$ be optimal covariance matrices for P1 and P2 respectively. We will need the following technical Lemma, which follows from von Neumann trace inequality [32].

Lemma 7: Let $\mathbf{A}$ *and* $\mathbf{B}$ *be positive semi-definite matrices and arrange their eigenvalues in decreasing order. Then,*

$$\mathrm{tr}(\mathbf{A}\mathbf{B}) \leq \sum_i \lambda_i(\mathbf{A})\lambda_i(\mathbf{B}). \quad (136)$$

Now define $\tilde{\mathbf{R}}' = \mathbf{U}_{30}^+\mathbf{R}^*\mathbf{U}_{30}$ and observe that

$$\begin{aligned} tr(\tilde{\mathbf{R}}') &= tr(\mathbf{U}_{30}\mathbf{U}_{30}^+\mathbf{R}^*) \\ &\leq \sum_i \lambda_i(\mathbf{U}_{30}\mathbf{U}_{30}^+)\lambda_i(\mathbf{R}^*) \\ &\leq \sum_i \lambda_i(\mathbf{R}^*) = tr\mathbf{R}^* \leq P_T \end{aligned} \quad (137)$$

where 1st inequality in (137) follows from Lemma 7; 2nd inequality is due to the fact that $\lambda_i(\mathbf{U}_{30}\mathbf{U}_{30}^+) = 0$ or 1, since $\mathbf{U}_{30}$ is semi-unitary. Thus, $\tilde{\mathbf{R}}'$ is feasible for P2 and hence

$$\begin{aligned} C_2 \geq \tilde{C}(\tilde{\mathbf{R}}') &= \ln \frac{|\mathbf{I} + \mathbf{U}_{30}^+\mathbf{W}_1\mathbf{U}_{30}\mathbf{U}_{30}^+\mathbf{R}^*\mathbf{U}_{30}|}{|\mathbf{I} + \mathbf{U}_{30}^+\mathbf{W}_2\mathbf{U}_{30}\mathbf{U}_{30}^+\mathbf{R}^*\mathbf{U}_{30}|} \\ &= \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{P}_3\mathbf{R}^*\mathbf{P}_3|}{|\mathbf{I} + \mathbf{W}_2\mathbf{P}_3\mathbf{R}^*\mathbf{P}_3|} \\ &= C(\mathbf{P}_3\mathbf{R}^*\mathbf{P}_3) = C(\mathbf{R}^*) = C_1 \end{aligned} \quad (138)$$

where

$$\tilde{C}(\tilde{\mathbf{R}}) = \ln|\mathbf{I} + \tilde{\mathbf{W}}_1\tilde{\mathbf{R}}| - \ln|\mathbf{I} + \tilde{\mathbf{W}}_2\tilde{\mathbf{R}}| \quad (139)$$

and $C(\mathbf{P}_3\mathbf{R}^*\mathbf{P}_3) = C(\mathbf{R}^*)$ follows from Lemma 2. Thus, $C_1 \leq C_2$.

Next, define $\mathbf{R}' = \mathbf{U}_{30}\tilde{\mathbf{R}}^*\mathbf{U}_{30}^+$ and observe that

$$tr(\tilde{\mathbf{R}}') = tr(\mathbf{U}_{30}^+\mathbf{U}_{30}\tilde{\mathbf{R}}^*) = tr\tilde{\mathbf{R}}^* \leq P_T \quad (140)$$

where 2nd equality follows from $\mathbf{U}_{30}^+\mathbf{U}_{30} = \mathbf{I}$. Additionally, $\mathbf{H}_3\mathbf{R}' = 0$ so that $\mathbf{R}'$ is feasible for P1 and hence

$$C_1 \geq C(\mathbf{R}') = \ln \frac{|\mathbf{I} + \mathbf{U}_{30}^+\mathbf{W}_1\mathbf{U}_{30}\tilde{\mathbf{R}}^*|}{|\mathbf{I} + \mathbf{U}_{30}^+\mathbf{W}_2\mathbf{U}_{30}\tilde{\mathbf{R}}^*|} = \tilde{C}(\tilde{\mathbf{R}}^*) = C_2$$

Combining this with $C_1 \leq C_2$, $C_1 = C_2$ follows. It is also clear that $\mathbf{R}^*$ and $\tilde{\mathbf{R}}^*$ are related as in (27).

G. Proof of Theorerm 2

Since the original problem P1 is equivalent to the projected problem P2, as was established in Proposition 4, we consider here P2. The corresponding Lagrangian is

$$\begin{aligned} L = \ln|\mathbf{I} + \tilde{\mathbf{W}}_2\tilde{\mathbf{R}}| - \ln|\mathbf{I} + \tilde{\mathbf{W}}_1\tilde{\mathbf{R}}| \\ + \lambda(tr(\tilde{\mathbf{R}}) - P_T) - tr(\mathbf{M}\tilde{\mathbf{R}}) \end{aligned} \quad (141)$$

where $\lambda \geq 0$ is the Lagrange multiplier responsible for the TPC $tr(\tilde{\mathbf{R}}) \leq P_T$ and $\mathbf{M} \geq \mathbf{0}$ is a matrix Lagrange multiplier responsible for $\tilde{\mathbf{R}} \geq \mathbf{0}$. The KKT conditions are:

$$(\tilde{\mathbf{W}}_2^{-1} + \tilde{\mathbf{R}})^{-1} - (\tilde{\mathbf{W}}_1^{-1} + \tilde{\mathbf{R}})^{-1} + \lambda\mathbf{I} - \mathbf{M} = 0, \quad (142)$$

$$\mathbf{M}\tilde{\mathbf{R}} = \mathbf{0}, \quad \lambda(tr(\tilde{\mathbf{R}}) - P_T) = 0, \quad (143)$$

$$\lambda \geq 0, \quad \mathbf{M} \geq \mathbf{0}, \ \tilde{\mathbf{R}} \geq \mathbf{0}, \ tr(\tilde{\mathbf{R}}) \geq 0. \quad (144)$$

Since $\tilde{\mathbf{W}}_1 - \tilde{\mathbf{W}}_2 > \mathbf{0}$, the objective in (25) is concave and hence the problem P2 is convex so that the KKT conditions (142)-(144) are sufficient for global optimality. It is straightforward to see that $\lambda > 0$ so that $tr(\tilde{\mathbf{R}}) = P_T$, i.e. the total power constraint is always active.

To establish (30), we first find $\tilde{\mathbf{R}}$ from the KKT conditions. Since $\tilde{\mathbf{R}} > 0$, then $\mathbf{M} = \mathbf{0}$ from (143) and (142) becomes

$$\mathbf{R}_1^{-1} - \mathbf{R}_2^{-1} = \lambda \mathbf{I} \tag{145}$$

where

$$\mathbf{R}_k = \tilde{\mathbf{W}}_k^{-1} + \tilde{\mathbf{R}}, \quad k = 1, 2. \tag{146}$$

so that $\mathbf{R}_1$ and $\mathbf{R}_2$ have same eigenvectors and their eigenvalue decomposition becomes $\mathbf{R}_k = \tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_k\tilde{\mathbf{U}}^+$, where the unitary matrix $\tilde{\mathbf{U}}$ collects the eigenvectors, the diagonal matrix $\tilde{\boldsymbol{\Lambda}}_k$ - corresponding eigenvalues. (145) can be expressed as

$$\lambda \mathbf{I} = (\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_1\tilde{\mathbf{U}}^+)^{-1} - (\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_2\tilde{\mathbf{U}}^+)^{-1} = \tilde{\boldsymbol{\Lambda}}_1^{-1} - \tilde{\boldsymbol{\Lambda}}_2^{-1}. \tag{147}$$

In addition,

$$\tilde{\mathbf{W}}_2^{-1} - \tilde{\mathbf{W}}_1^{-1} = \tilde{\mathbf{U}}(\tilde{\boldsymbol{\Lambda}}_2 - \tilde{\boldsymbol{\Lambda}}_1)\tilde{\mathbf{U}}^+ \tag{148}$$

so that the columns of $\tilde{\mathbf{U}}$ are also the eigenvectors of $\tilde{\mathbf{W}}_2^{-1} - \tilde{\mathbf{W}}_1^{-1}$, and hence from (146)

$$\tilde{\mathbf{R}}^* = \tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_1\tilde{\mathbf{U}}^+ - \tilde{\mathbf{W}}_1^{-1} \tag{149}$$

where $\tilde{\boldsymbol{\Lambda}}_1$ is found from (147) and (148), as expressed in (31). The unprojected covariance $\mathbf{R}^*$ in (30) follows from (149). (32) follows from $tr\mathbf{R} = tr(\tilde{\mathbf{R}}) = P_T$ and (149).

To establish the threshold power P_{T0} in (29), observe that $\tilde{\mathbf{R}}^* > \mathbf{0}$ iff $\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_1\tilde{\mathbf{U}}^+ > \tilde{\mathbf{W}}_1^{-1}$. It follows from (31) that λ_{1i} are decreasing functions of λ so that, as P_T increases, λ decreases and hence all λ_{1i} increase. Furthermore, $\tilde{\mathbf{U}}\tilde{\boldsymbol{\Lambda}}_1\tilde{\mathbf{U}}^+ > \tilde{\mathbf{W}}_1^{-1}$ if $\min_i \lambda_{1i}\lambda_{min} > 1$ and P_{T0} is found from the boundary condition $\min_i \lambda_{1i}\lambda_{min} = 1$, which implies

$$\frac{1}{\lambda_{min}} = \frac{2}{\lambda}\left(\sqrt{1 + \frac{4\mu_1}{\lambda}} + 1\right)^{-1} \tag{150}$$

from which (29) follows after some manipulations.

To prove $r(\mathbf{R}^*) = m - r_3$, notice that

$$\lambda_i(\mathbf{R}^*) = \lambda_i(\tilde{\mathbf{U}}_{3-}^+\tilde{\mathbf{U}}_{3-}\tilde{\mathbf{R}}^*) = \lambda_i(\tilde{\mathbf{R}}^*) \tag{151}$$

so that $\tilde{\mathbf{R}}^*$ and $\mathbf{R}^*$ have the same (non-zero) eigenvalues and hence the same rank, and $r(\tilde{\mathbf{R}}^*) = m - r_3$ (since it is full-rank). Finally, (30) follows from (27) and (32) - from the TPC.

H. Proof of Proposition 5

To establish 1st inequality, use Proposition 4 and consider problem P2 instead, whose optimal covariance is $\tilde{\mathbf{R}}^* = \tilde{\mathbf{U}}_+\widetilde{\boldsymbol{\Lambda}_+}\tilde{\mathbf{U}}_+^+$, where $\tilde{\mathbf{U}}_+$ is the semi-unitary matrix of its active eigenvectors and $\tilde{\boldsymbol{\Lambda}}_+$ is the diagonal matrix of its positive eigenvalues. It follows from Proposition 3 in [24] applied to P2 that

$$\tilde{\mathbf{U}}_+^+(\tilde{\mathbf{W}}_1 - \tilde{\mathbf{W}}_2)\tilde{\mathbf{U}}_+ = \tilde{\mathbf{U}}_+^+\mathbf{U}_{30}^+(\mathbf{W}_1 - \mathbf{W}_2)\mathbf{U}_{30}\tilde{\mathbf{U}}_+ > \mathbf{0}$$

and, from (27), that

$$\mathbf{R}^* = \mathbf{U}_{30}\tilde{\mathbf{U}}_+\tilde{\boldsymbol{\Lambda}}_+\tilde{\mathbf{U}}_+^+\mathbf{U}_{30}^+ = \mathbf{U}_+\tilde{\boldsymbol{\Lambda}}_+\mathbf{U}_+^+ \tag{152}$$

where $\mathbf{U}_+ = \mathbf{U}_{30}\tilde{\mathbf{U}}_+$ is a semi-unitary matrix of active eigenvectors of $\mathbf{R}^*$, which establishes 1st inequality. 2nd one follows from 1st one.

I. Proof of Corollary 1

To establish this result, observe from Proposition 6 that, under stated condition, $r(\mathbf{R}^*) = r(\tilde{\mathbf{R}}^*) = 1$. Then, use the problem equivalence in Proposition 4 and apply rank-1 solution in [24] to P2 so that $\tilde{\mathbf{R}}^* = P_T\mathbf{u}_1\mathbf{u}_1^+$, from which the desired result follows.

J. Proof of Proposition 7

Note that $\mathcal{R}(\mathbf{W}_2) \in \mathcal{R}(\mathbf{W}_3)$ and $\mathbf{R}^*\mathbf{W}_3 = 0$ imply $\mathbf{R}^*\mathbf{W}_2 = 0$, i.e. $\mathcal{R}(\mathbf{R}^*) \in \mathcal{N}(\mathbf{W}_2) \cap \mathcal{N}(\mathbf{W}_3) = \mathcal{N}(\mathbf{W}_3)$ and

$$\begin{aligned} C(\mathbf{R}^*) &= \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}^*| \\ &= \max_{\mathbf{R}\geq 0} \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| \text{ s.t. } tr\mathbf{R} \leq P_T, \ \mathbf{W}_3\mathbf{R} = 0 \\ &= \max_{\tilde{\mathbf{R}}\geq 0} \ln|\mathbf{I} + \tilde{\mathbf{W}}_1\tilde{\mathbf{R}}| \text{ s.t. } tr\tilde{\mathbf{R}} \leq P_T, \end{aligned} \tag{153}$$

for which the solution is given by the water-filling over the eigenmodes of $\tilde{\mathbf{W}}_1$, $\tilde{\mathbf{R}}^* = (\mu^{-1}\mathbf{I} - \tilde{\mathbf{W}}_1^{-1})_+$, which, when transformed back to the original space, gives (35).

K. Proof of Proposition 8

First, observe that adding constraints cannot increase the capacity, so that $C_{s1} \geq C_s$, where C_{s1} and C_s are the single-Ev single-PR and multi-Ev multi-PR capacities, respectively. Let $\mathbf{R}_1^*$ be an optimal covariance for the single-user case, so that

$$P_I \geq tr\mathbf{W}_{31}\mathbf{R}_1^* \geq tr\mathbf{W}_{3k}\mathbf{R}_1^* \quad \forall k \tag{154}$$

where 2nd inequality is due to the ordering $\mathbf{W}_{31} \geq \mathbf{W}_{3k}$, and hence $\mathbf{R}_1^*$ is also feasible for the multi-user case. Furthermore, the information leakage to k-th Ev does not exceed that to 1st one,

$$I(w, \mathbf{Y}_{21}^n) \geq I(w, \mathbf{Y}_{2k}^n) \tag{155}$$

since $\mathbf{W}_{21} \geq \mathbf{W}_{2k}$, and, hence, if the secrecy criterion is satisfied for 1st Ev, so it is for k-th Ev, i.e. any wiretap code that works in the single-user case, also works in the multi-user case, and hence $C_{s1} = C_s$.

L. Proof of Proposition 9

To prove sufficiency, observe that (40) implies that $\exists \mathbf{u}$: $\mathbf{u}^+\mathbf{W}_2\mathbf{u} = 0$, $\mathbf{u}^+\mathbf{W}_1\mathbf{u} > 0$, so that one can set $\mathbf{R} = P_T\mathbf{u}\mathbf{u}^+$ and

$$C_s \geq C(\mathbf{R}) = \ln(1 + P_T\mathbf{u}^+\mathbf{W}_1\mathbf{u}) \rightarrow \infty \tag{156}$$

as $P_T \rightarrow \infty$.

The proof of necessity is by contradiction: assume that $\mathcal{N}(\mathbf{W}_2) \in \mathcal{N}(\mathbf{W}_1)$, which implies $\mathcal{R}(\mathbf{W}_1) \in \mathcal{R}(\mathbf{W}_2)$ and hence $\mathbf{P}_2\mathbf{W}_k\mathbf{P}_2 = \mathbf{W}_k$, where $\mathbf{P}_2 = \mathbf{U}_{2+}\mathbf{U}_{2+}^+$ is a projection matrix on $\mathcal{R}(\mathbf{W}_2)$ and $\mathbf{U}_{2+}$ is a semi-unitary matrix of active eigenvectors of $\mathbf{W}_2$. Let us re-normalize $\mathbf{R}$ as follows:

$\mathbf{R} \rightarrow \mathbf{R}/P_T$ so that $tr\mathbf{R} \leq 1$ and the actual Tx covariance is $P_T\mathbf{R}$. It follows that

$$C(\mathbf{R}) = \ln\frac{|\mathbf{I}+P_T\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+P_T\mathbf{W}_2\mathbf{R}|} = \ln\frac{|\mathbf{I}+P_T\mathbf{P}_2\mathbf{W}_1\mathbf{P}_2\mathbf{R}|}{|\mathbf{I}+P_T\mathbf{P}_2\mathbf{W}_2\mathbf{P}_2\mathbf{R}|} = \ln\frac{|\mathbf{I}+P_T\mathbf{W}_1'\mathbf{R}'|}{|\mathbf{I}+P_T\mathbf{W}_2'\mathbf{R}'|} \quad (157)$$

where $\mathbf{W}_k' = \mathbf{U}_{2+}^{+}\mathbf{W}_k\mathbf{U}_{2+}$, $\mathbf{R}' = \mathbf{U}_{2+}^{+}\mathbf{R}\mathbf{U}_{2+}$, and $tr\mathbf{R}' \leq tr\mathbf{R} \leq 1$. Note that $\mathbf{W}_2' > 0$ and hence

$$C(\mathbf{R}) = \sum_{i=1}^{r_2}\ln\frac{1+P_T\lambda_i(\mathbf{W}_1'\mathbf{R}')}{1+P_T\lambda_i(\mathbf{W}_2'\mathbf{R}')} \leq \sum_{i=1}^{r_2}\ln\frac{1+P_T\lambda_1(\mathbf{W}_1')\lambda_i(\mathbf{R}')}{1+P_T\lambda_{r2}(\mathbf{W}_2')\lambda_i(\mathbf{R}')} \leq r_2\ln\frac{\lambda_1(\mathbf{W}_1)}{\lambda_{r2}(\mathbf{W}_2)} < \infty \quad (158)$$

for any $\mathbf{R}$ and any P_T, where $r_2 = r(\mathbf{W}_2)$, $\lambda_{r2}(\mathbf{W}_2) > 0$. 1st inequality follows from the matrix eigenvalue inequalities: $\lambda_i(\mathbf{WR}) \leq \lambda_1(\mathbf{W})\lambda_i(\mathbf{R})$ and $\lambda_i(\mathbf{WR}) \geq \lambda_m(\mathbf{W})\lambda_i(\mathbf{R})$ [32]. 2nd inequality follows from the fact that the 2nd ratio in (158) is increasing in P_T. Hence,

$$C_s = \max_{\mathbf{R}} C(\mathbf{R}) \leq r_2\ln\frac{\lambda_1(\mathbf{W}_1)}{\lambda_{r2}(\mathbf{W}_2)} < \infty \quad (159)$$

M. Proof of Proposition 10

Sufficiency can be established in the same way as in Proposition 9: (41) implies that $\exists\mathbf{z} : \mathbf{z}^+\mathbf{W}_2\mathbf{z} = \mathbf{z}^+\mathbf{W}_3\mathbf{z} = 0$ while $\mathbf{z}^+\mathbf{W}_1\mathbf{z} > 0$ so that setting $\mathbf{R} = P_T\mathbf{z}\mathbf{z}^+$ (which is feasible), one obtains:

$$C_s \geq C(\mathbf{R}) = \ln(1+P_T\mathbf{z}^+\mathbf{W}_1\mathbf{z}) > 0 \quad (160)$$

To establish the necessary part, let us use the normalized covariance $tr\mathbf{R} \leq 1$ (where the actual Tx covariance is $P_T\mathbf{R}$). It follows from the IPC that $P_T tr\mathbf{W}_3\mathbf{R} \leq P_I$ and hence

$$\lambda_i(\mathbf{R})\mathbf{u}_i^+\mathbf{W}_3\mathbf{u}_i \leq P_I/P_T \rightarrow 0 \quad (161)$$

as $P_T \rightarrow \infty$ under bounded P_I, which implies that all active eigenvectors $\mathbf{u}_{i+}$ of $\mathbf{R}$ satisfy $\mathbf{u}_{i+}^+\mathbf{W}_3\mathbf{u}_{i+} = 0$ asymptotically, i.e. $\mathbf{u}_{i+} \in \mathcal{N}(\mathbf{W}_3)$. Hence, $\mathbf{R} = \mathbf{P}_3\mathbf{R}\mathbf{P}_3$, where $\mathbf{P}_3 = \mathbf{U}_{30}\mathbf{U}_{30}^+$ is a projection matrix on $\mathcal{N}(\mathbf{W}_3)$ and $\mathbf{U}_{30}$ is a semi-unitary matrix of inactive eigenvectors of $\mathbf{W}_3$, which also form an orthonormal basis of $\mathcal{N}(\mathbf{W}_3)$. With this in mind, $C(\mathbf{R})$ can be presented as

$$C(\mathbf{R}) = \ln\frac{|\mathbf{I}+P_T\mathbf{W}_1\mathbf{R}|}{|\mathbf{I}+P_T\mathbf{W}_2\mathbf{R}|} = \ln\frac{|\mathbf{I}+P_T\mathbf{W}_1\mathbf{P}_3\mathbf{R}\mathbf{P}_3|}{|\mathbf{I}+P_T\mathbf{W}_2\mathbf{P}_3\mathbf{R}\mathbf{P}_3|} = \ln\frac{|\mathbf{I}+P_T\mathbf{W}_1'\mathbf{R}'|}{|\mathbf{I}+P_T\mathbf{W}_2'\mathbf{R}'|} \quad (162)$$

where $\mathbf{W}_k' = \mathbf{U}_{30}^+\mathbf{W}_k\mathbf{U}_{30}$, $\mathbf{R}' = \mathbf{U}_{30}^+\mathbf{R}\mathbf{U}_{30}$, and the secrecy capacity is

$$C_s = \max_{\mathbf{R}' \geq 0} C(\mathbf{R}') \quad \text{s.t. } tr\mathbf{R}' \leq P_T \quad (163)$$

Now one can use the condition in (40) with $\mathbf{W}_k'$ in place of $\mathbf{W}_k$ to conclude that the capacity saturates iff:

$$\mathcal{N}(\mathbf{W}_2') \in \mathcal{N}(\mathbf{W}_1') \quad (164)$$

i.e. if

$$\mathbf{z}^+\mathbf{W}_2'\mathbf{z} = \mathbf{z}^+\mathbf{U}_{30}^+\mathbf{W}_2\mathbf{U}_{30}\mathbf{z} = 0 \quad (165)$$

implies

$$\mathbf{z}^+\mathbf{W}_1'\mathbf{z} = \mathbf{z}^+\mathbf{U}_{30}^+\mathbf{W}_1\mathbf{U}_{30}\mathbf{z} = 0 \quad (166)$$

for any $\mathbf{z}$. The last equality in (165) implies that $\mathbf{z}' = \mathbf{U}_{30}\mathbf{z} \in \mathcal{N}(\mathbf{W}_3) \cap \mathcal{N}(\mathbf{W}_2)$ and likewise (166) implies that $\mathbf{z}' \in \mathcal{N}(\mathbf{W}_3) \cap \mathcal{N}(\mathbf{W}_1)$ so that (164) is equivalent to

$$\mathcal{N}(\mathbf{W}_3) \cap \mathcal{N}(\mathbf{W}_2) \in \mathcal{N}(\mathbf{W}_3) \cap \mathcal{N}(\mathbf{W}_1) \quad (167)$$

which is in turn equivalent to $\mathcal{N}(\mathbf{W}_3) \cap \mathcal{N}(\mathbf{W}_2) \in \mathcal{N}(\mathbf{W}_1)$ so that the capacity grows unbounded iff (41) holds.

Acknowledgement

The authors would like to thank the reviewers for their constructive comments and suggestions.

References

[1] M. Shafi *et al.*, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201–1221, Jun. 2017.

[2] V. W. S. Wong Ed. *Key Technologies for 5G Wireless Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[3] W. Nam, D. Bai, J. Lee, and I. Kang, "Advanced interference management for 5G cellular networks," *IEEE Commun. Mag.*, vol. 52, no. 5, pp. 52–60, May 2014.

[4] L. Sanguinetti, A. L. Moustakas, and M. Debbah, "Interference management in 5G reverse TDD HetNets with wireless backhaul: A large system analysis," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 33, no. 6, pp. 1187–1200, Jun. 2015.

[5] H. Zhang, Y. Liao, and L. Song, "D2D-U: Device-to-device communications in unlicensed bands for 5G system," *IEEE Trans. Commun.*, vol. 16, no. 6, pp. 3507–3519, Jun. 2017.

[6] L. Song, Y. Li, Z. Ding, and H. V. Poor, "Resource management in non-orthogonal multiple access networks for 5G and beyond," *IEEE Netw.*, vol. 31, no. 4, pp. 8–14, Jul./Aug. 2017.

[7] W. Liang, "Non-orthogonal multple access (NOMA) for 5G systems," in *Key Technologies for 5G Wireless Systems*, V. W. S. Wong, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[8] S. Haykin *et al.*, "Cognitive radio, part 1: Practical perspectives, and part 2: Fundamental issues," *Proc. IEEE*, vol. 97, nos. 4–5, Apr./May 2009.

[9] K. Haghighi, E. G. Strom, and E. Agrell, "On optimum causal cognitive spectrum reutilization strategy," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 10, pp. 1911–1921, Nov. 2012.

[10] X. Hong, J. Wang, C.-X. Wang, and J. Shi, "Cognitive radio in 5G: A perspective on energy-spectral efficiency trade-off," *IEEE Commun. Mag.*, vol. 52, no. 7, pp. 46–53, May 2014.

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

[12] Q. Zhang, S. Kota, V. Lau, W. Su, and A. Kwasinski, "Introduction to the issue on cooperative communication and signal processing in cognitive radio systems," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 1, pp. 1–4, Feb. 2011.

[13] R. Zhang and Y.-C. Liang, "Exploiting multi-antennas for opportunistic spectrum sharing in cognitive radio networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 88–102, Feb. 2008.

[14] A. G. Fragkiadakis, E. Z. Tragos, and I. G. Askoxylakis, "A survey on security threats and detection techniques in cognitive radio networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 428–445, 1st Quart., 2013.

[15] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[16] P. A. Regalia, A. Khisti, Y. Liang, and S. Tomasin, "Secure communications via physical-layer and information-theoretic techniques," *Proc. IEEE*, vol. 103, no. 10, pp. 1698–1701, Oct. 2015.

[17] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas I: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3088–3104, Jul. 2010.

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

[19] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[20] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory.*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[21] J. Li and A. Petropulu. (Sep. 2009). "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels." [Online]. Available: https://arxiv.org/abs/0909.2622

[22] J. Li *et al.*, "Multiple antennas for physical layer secrecy," in *Trends in Digital Signal Processing*, Y. C. Lim, Ed. Singapore: Pan Stanford, 2015.

[23] S. Loyka and C. D. Charalambous, "Rank-deficient solutions for optimal signaling over wiretap MIMO channels," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2400–2411, Jun. 2016.

[24] S. Loyka and C. D. Charalambous, "Optimal signaling for secure communications over Gaussian MIMO wiretap channels," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7207–7215, Dec. 2016.

[25] S. Loyka and C. D. Charalambous, "An algorithm for global maximization of secrecy rates in Gaussian MIMO wiretap channels," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2288–2299, Jun. 2015.

[26] L. Zhang, R. Zhang, Y.-C. Liang, Y. Xin, and S. Cui, "On the relationship between the multi-antenna secrecy communications and cognitive radio communications," *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1877–1886, Jun. 2010.

[27] Y. Pei, Y.-C. Liang, L. Zhang, K. C. Teh, and K. H. Li, "Secure communication over MISO cognitive radio channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1494–1502, Apr. 2010.

[28] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[29] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory.*, vol. IT-33, no. 5, pp. 711–716, Sep. 1987.

[30] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[31] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Belmont, MA, USA: Athena Scientific, 2008.

[32] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 2013.

[33] L. Dong, S. Loyka, and Y. Li, "The operational secrecy capacity of cognitive radio MIMO channel," in *Proc. 15th Can. Workshop Inf. Theory*, Quebec City, QC, Canada, Jun. 2017, pp. 1–5.

[34] L. Dong, S. Loyka, and Y. Li, "Closed-form solutions for optimal secure signaling over cognitive radio MIMO channels (Invited Paper)," in *Proc. 5th IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Montreal, QC, Canada, Nov. 2017, pp. 502–506.

[35] Q. Li and W.-K. Ma, "Spatially selective artificial-noise aided transmit optimization for MISO multi-eves secrecy rate maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2704–2717, May 2013.

[36] Y. Pei, Y.-C. Liang, K. C. Teh, and K. H. Li, "Secure communication in multiantenna cognitive radio networks with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1683–1693, Apr. 2011.

[37] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound wiretap channels," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, Oct. 2009, Art. no. 142374.

[38] R. F. Schaefer and S. Loyka, "The secrecy capacity of compound Gaussian MIMO wiretap channels," *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5535–5552, Oct. 2015.

[39] S. Loyka and L. Dong, "Optimal full-rank signalling over MIMO wiretap channels under interference constraint," *IEEE Wireless Commun. Lett.*, to be published.

Limeng Dong was born in Xi'an, China. He received the bachelor's and master's degrees from the School of Electronics and Information, Northwestern Polytechnical University, China, where he is currently pursuing the Ph.D. degree. From 2015 to 2017, he was a Visiting Ph.D. Student with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. His research interests include wireless communications, cognitive radio, and physical-layer security.

Sergey Loyka was born in Minsk, Belarus. He received the M.S. degree (Hons.) from the Minsk Radioengineering Institute, Minsk, in 1992, and the Ph.D. degree in radio engineering from the Belorussian State University of Informatics and Radioelectronics (BSUIR), Minsk, in 1995. He was a Research Fellow with the Laboratory of Communications and Integrated Microelectronics, École de Technologie Supérieure, Montreal, Canada, a Senior Scientist with the Electromagnetic Compatibility Laboratory, BSUIR, an invited Scientist with the Laboratory of Electromagnetism and Acoustic, Swiss Federal Institute of Technology, Lausanne, Switzerland. Since 2001, he has been a Faculty Member with the School of Electrical Engineering and Computer Science, University of Ottawa, Canada. His research interests include wireless communications and networks and, in particular, MIMO systems and the security aspects of such systems, in which he has published extensively. He received a number of awards from the URSI, the IEEE, the Swiss, Belarus and former USSR governments, and the Soros Foundation.

Yong Li received the Ph.D. degree from Northwestern Polytechnical University, China. He is currently a Professor with the School of Electronics and Information, Northwestern Polytechnical University. His research interests include digital signal processing with its applications, the design and development of high speed real-time DSP systems, and the cognitive radio and the software-defined radio technologies.