Full Rank Solutions for the MIMO Gaussian Wiretap Channel with an Average Power Constraint

S. Ali. A. Fakoorian, *Student Member, IEEE* and A. Lee Swindlehurst, *Fellow, IEEE*

Abstract

This paper considers a multiple-input multiple-output (MIMO) Gaussian wiretap channel model, where there exists a transmitter, a legitimate receiver and an eavesdropper, each equipped with multiple antennas. In this paper, we first revisit the rank property of the optimal input covariance matrix that achieves the secrecy capacity of the multiple antenna MIMO Gaussian wiretap channel under the average power constraint. Next, we obtain necessary and sufficient conditions on the MIMO wiretap channel parameters such that the optimal input covariance matrix is full-rank, and we fully characterize the resulting covariance matrix as well. Numerical results are presented to illustrate the proposed theoretical findings.

Index Terms

MIMO Wiretap Channel, Secrecy Capacity, Physical Layer Security

EDICS: WIN-PHYL, WIN-INFO, MSP-CAPC, WIN-CONT

The authors are with the Dept. of Electrical Engineering and Computer Science, University of California, Irvine, CA 92697-2625, USA. e-mail:{afakoori, swindle}@uci.edu

I. INTRODUCTION

The broadcast nature of a wireless medium makes it very susceptible to eavesdropping, where the transmitted message is decoded by unintended receiver(s). Recent information-theoretic research on secure communication has focused on enhancing security at the physical layer. The wiretap channel, first introduced and studied by Wyner [1], is the most basic physical layer model that captures the problem of communication security. Wyner showed that when an eavesdropper's channel is a degraded version of the main channel, the source and destination can achieve a positive secrecy rate, while ensuring that the eavesdropper gets zero bits of information. The maximum secrecy rate from the source to the destination is defined as the secrecy capacity. The Gaussian wiretap channel, in which the outputs at the legitimate receiver and at the eavesdropper are corrupted by additive white Gaussian noise, was studied in [2].

Determining the secrecy capacity of a Gaussian wiretap channel is in general a difficult non-convex optimization problem, and has been addressed independently in [3]-[9]. Oggier and Hassibi [3] and Khisti and Wornell [4] followed an indirect approach using a Sato-like argument and matrix analysis tools. They considered the problem of finding the secrecy capacity of the Gaussian MIMO wiretap channel under the average total power constraint, and a closed-form expression for the secrecy capacity in the high signal-to-noise-ratio (SNR) regime was obtained in [4]. In [5], the rank property of the optimal input covariance matrix for the secrecy rate maximization problem is discussed but the authors were unable to characterize the solution for the general case. For some special cases of the MIMO wiretap channel, where the solution has rank one, the optimal input covariance matrix that achieves the secrecy capacity under the average total power constraint was obtained in [5]-[7].

In [8], Liu and Shamai propose a more information-theoretic approach using the enhancement concept, originally presented by Weingarten et al. [10], as a tool for the characterization of the MIMO Gaussian broadcast channel capacity. Liu and Shamai have shown that an enhanced degraded version of the channel attains the same secrecy capacity as does a Gaussian input distribution. From the mathematical solution in [8] it was evident that such an enhanced channel exists; however it was not clear how to construct such a channel until the work of [9], which provided a closed-form expression for the secrecy capacity under a *covariance matrix* power constraint. While this result is interesting since the expression for the secrecy capacity is valid for all SNR scenarios, there still exists no computable secrecy capacity expression for the MIMO Gaussian wiretap channel under an average total power constraint.

In this paper, we first investigate the rank property of the optimal input covariance matrix that achieves

the secrecy capacity of the general Gaussian multiple-input multiple-output (MIMO) wiretap channel under the average total power constraint, where the number of antennas is arbitrary for both the transmitter and the two receivers. Next, we obtain the optimal input covariance matrix for the case that this optimal covariance matrix is full-rank. Necessary and sufficient conditions to have a full-rank optimal input covariance matrix are characterized as well.

The rest of this paper is organized as follows. In the next section, we describe the assumed mathematical model and revisit the current solution for the wiretap channel under the matrix power constraint. The rank property of the optimal input covariance matrix under the average power constraint is investigated in Section III, and in Section IV we characterize the conditions under which the input covariance matrix that achieves the secrecy capacity of a wiretap channel under the average power constraint is full-rank. In Section V, we discuss some interesting facts regarding the optimal solution, and in Section VI we present numerical results to illustrate the proposed solutions. Finally, Section VII concludes the paper.

Notation: Vector-valued random variables are written with non-boldface uppercase letters (*e.g.,* X), while the corresponding non-boldface lowercase letter ($\mathbf{x}$) denotes a specific realization of the random variable. Scalar variables are written with non-boldface (lowercase or uppercase) letters. The Hermtian (i.e., conjugate) transpose is denoted by $(.)^H$, the matrix trace by Tr$(.)$, and $\mathbf{I}$ indicates an identity matrix. Inequality $\mathbf{A} \preceq \mathbf{B}$ means that $\mathbf{A} - \mathbf{B}$ is Hermitian positive semi-definite. The Euclidean norm of the vector $\mathbf{x}$ is written as $\|\mathbf{x}\|$. Mutual information between the random variables A and B is denoted by $I(A; B)$, E is the expectation operator, and $\mathcal{CN}(0, \sigma^2)$ represents the complex circularly symmetric Gaussian distribution with zero mean and variance σ^2.

II. SYSTEM MODEL AND PRIOR WORKS

We begin with a multiple-antenna wiretap channel with n_t transmit antennas and n_r and n_e receive antennas at the legitimate recipient and the eavesdropper, respectively:

$$\mathbf{y}_r = \mathbf{H}\mathbf{x} + \mathbf{z}_r$$
$$\mathbf{y}_e = \mathbf{G}\mathbf{x} + \mathbf{z}_e$$

$$(1)$$

where $\mathbf{x}$ is a zero-mean $n_t \times 1$ transmitted signal vector, $\mathbf{z}_r \in \mathbb{C}^{n_r \times 1}$ and $\mathbf{z}_e \in \mathbb{C}^{n_e \times 1}$ are additive white Gaussian noise vectors at the receiver and eavesdropper, respectively, with i.i.d. entries distributed as $\mathcal{CN}(0, 1)$. The matrices $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ and $\mathbf{G} \in \mathbb{C}^{n_e \times n_t}$ represent the channels associated with the receiver and the eavesdropper, respectively. Similar to other papers considering the perfect secrecy rate of the wiretap channel, we assume that the transmitter has perfect channel state information (CSI) for

both the legitimate receiver and the eavesdropper. For the Gaussian channel, where Gaussian inputs are an optimal choice, the secrecy capacity is given by [3]

$$\mathcal{C}_{sec} = \max_{\mathbf{x}}[I(X;Y_r) - I(X;Y_e)] = \max_{\mathbf{Q} \succeq \underline{0}} R(\mathbf{Q}) \tag{2}$$

where $R(\mathbf{Q}) = \log|\mathbf{HQH}^H + \mathbf{I}| - \log|\mathbf{GQG}^H + \mathbf{I}|$, and $\mathbf{Q} = E\{\mathbf{xx}^H\}$ is the input covariance matrix.

In [9], the above secret communication problem was analyzed under the matrix power-covariance constraint, defined as

$$\mathbf{Q} \preceq \mathbf{S} \tag{3}$$

where $\mathbf{S}$ is a positive semi-definite matrix. An explicit expression for the secrecy capacity under (3) was obtained via applying the generalized eigenvalue decomposition to the following two positive definite matrices

$$(\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{HS}^{\frac{1}{2}} + \mathbf{I} \quad , \quad \mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{GS}^{\frac{1}{2}} + \mathbf{I}) \tag{4}$$

In particular, there exists an invertible generalized eigenvector matrix $\mathbf{C}$ such that [14]

$$\mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{GS}^{\frac{1}{2}} + \mathbf{I}\right] \mathbf{C} = \mathbf{I} \tag{5}$$

$$\mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{HS}^{\frac{1}{2}} + \mathbf{I}\right] \mathbf{C} = \mathbf{\Lambda} \tag{6}$$

where $\mathbf{\Lambda} = \text{diag}\{\lambda_1, ..., \lambda_{n_t}\}$ is a positive definite diagonal matrix and $\lambda_1, ..., \lambda_{n_t}$ represent the generalized eigenvalues. Without loss of generality, we assume the eigenvalues are ordered as

$$\lambda_1 \geq ... \geq \lambda_b > 1 \geq \lambda_{b+1} \geq ... \geq \lambda_{n_t} > 0$$

so that a total of b $(0 \leq b \leq n_t)$ are greater than 1. Hence, we can write $\mathbf{\Lambda}$ as

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & \underline{0} \\ \underline{0} & \mathbf{\Lambda}_2 \end{bmatrix} \tag{7}$$

where $\mathbf{\Lambda}_1 = \text{diag}\{\lambda_1, ..., \lambda_b\}$ and $\mathbf{\Lambda}_2 = \text{diag}\{\lambda_{b+1}, ..., \lambda_{n_t}\}$. We can partition $\mathbf{C}$ similarly:

$$\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2] \tag{8}$$

where $\mathbf{C}_1$ is the $n_t \times b$ submatrix representing the generalized eigenvectors corresponding to $\{\lambda_1, ..., \lambda_b\}$ and $\mathbf{C}_2$ is the $n_t \times (n_t - b)$ submatrix representing the generalized eigenvectors corresponding to $\{\lambda_{b+1}, ..., \lambda_{n_t}\}$. Using the above notation, the secrecy capacity of the MIMO wiretap channel under the *matrix* power constraint (3) can be expressed as [9], [13, Theorem 3]:

Corollary 1. Under the matrix power constraint (3), the secrecy capacity of the MIMO Gaussian wiretap channel is given by

$$\mathcal{C}_{sec}(\mathbf{S}) = \sum_{i=1}^{b} \log \lambda_i = \log |\mathbf{\Lambda}_1| \tag{9}$$

where the optimal input covariance matrix $\mathbf{Q}_S^*$ that maximizes (2) and attains (9) is given by

$$\mathbf{Q}_S^* = \mathbf{S}^{\frac{1}{2}} \mathbf{C} \begin{bmatrix} (\mathbf{C}_1^H \mathbf{C}_1)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{C}^H \mathbf{S}^{\frac{1}{2}} . \tag{10}$$

Remark 1. From (5) and (6), one can easily confirm that if $\mathbf{H}^H \mathbf{H} \preceq \mathbf{G}^H \mathbf{G}$, then for any $\mathbf{S} \succeq \underline{0}$ we have $\mathbf{\Lambda} \preceq \mathbf{I}$. In other words, in this case the pencil in (4) has no generalized eigenvalue bigger than 1. Thus, $\mathcal{C}_{sec}(\mathbf{S}) = 0$ for any $\mathbf{S} \succeq \underline{0}$.

In this paper, we consider the secrecy capacity problem in (2) under the *average* power constraint:

$$\text{Tr}(E\{\mathbf{x}\mathbf{x}^H\}) = \text{Tr}(\mathbf{Q}) \le P_t. \tag{11}$$

For this constraint, no computable secrecy capacity expression has been derived to date for the general MIMO case. In principle, one would have to find the secrecy capacity through an exhaustive search over the set $\{\mathbf{S} : \mathbf{S} \succeq \underline{0}, \text{Tr}(\mathbf{S}) \le P\}$ [10, Lemma 1], [13]:

$$\mathcal{C}_{sec}(P_t) = \max_{\mathbf{S} \succeq \underline{0}, \text{Tr}(\mathbf{S}) = P_t} \mathcal{C}_{sec}(\mathbf{S}) . \tag{12}$$

where for any given semidefinite $\mathbf{S}$, $\mathcal{C}_{sec}(\mathbf{S})$ should be computed as given by (9).

In the next section, we investigate the rank of the optimal input covariance matrix $\mathbf{Q}^*$ that attains $\mathcal{C}_{sec}(P_t)$. Next, in Section IV, we obtain the optimal $\mathbf{Q}^*$ under the average power constraint for the case that $\mathbf{Q}^*$ is full-rank.

III. RANK PROPERTY OF THE OPTIMAL SOLUTION UNDER AN AVERAGE POWER CONSTRAINT

First, we note that the problem under $\text{Tr}(\mathbf{Q}) \le P_t$ is equivalent to that under $\text{Tr}(\mathbf{Q}) = P_t$ [3], [5][1]. Also note that in (12), this implies that we have $\text{Tr}(\mathbf{S}) = P_t$ instead of $\text{Tr}(\mathbf{S}) \le P_t$.

We are interested in finding the optimal $\widehat{\mathbf{S}}$ which maximizes the problem (12). Let us assume that we have found the optimal $\widehat{\mathbf{S}}$. Consequently, from (10), the optimal input covariance matrix that attains

---

[1]For this statement, and also for the following results in the paper, we exclude the special case $\mathbf{H}^H \mathbf{H} \preceq \mathbf{G}^H \mathbf{G}$ for which the $\mathcal{C}_{sec}$ is trivially 0 for any $\mathbf{S} \succeq \underline{0}$, and consequently for any P_t, as pointed out in Remark 1.

$\mathcal{C}_{sec}(P_t)$ is given by

$$\mathbf{Q}^* = \widehat{\mathbf{S}}^{\frac{1}{2}}\widehat{\mathbf{C}} \begin{bmatrix} (\widehat{\mathbf{C}}_1^H\widehat{\mathbf{C}}_1)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \widehat{\mathbf{C}}^H\widehat{\mathbf{S}}^{\frac{1}{2}} , \tag{13}$$

where $\widehat{\mathbf{C}}$ and $\widehat{\mathbf{C}}_1$ have respectively the same definitions as those of $\mathbf{C}$ and $\mathbf{C}_1$, given by (5)-(8), but here for the pencil $(\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I}\, ,\ \widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I})$. Note that $\mathbf{Q}^*$ can be rewritten as

$$\mathbf{Q}^* = \widehat{\mathbf{S}}^{\frac{1}{2}} \begin{bmatrix} \widehat{\mathbf{C}}_1 & \widehat{\mathbf{C}}_2 \end{bmatrix} \begin{bmatrix} (\widehat{\mathbf{C}}_1^H\widehat{\mathbf{C}}_1)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} \widehat{\mathbf{C}}_1^H \\ \widehat{\mathbf{C}}_2^H \end{bmatrix} \widehat{\mathbf{S}}^{\frac{1}{2}}$$

$$= \widehat{\mathbf{S}}^{\frac{1}{2}}\, \widehat{\mathbf{C}}_1(\widehat{\mathbf{C}}_1^H\widehat{\mathbf{C}}_1)^{-1}\widehat{\mathbf{C}}_1^H\, \widehat{\mathbf{S}}^{\frac{1}{2}}$$

$$= \widehat{\mathbf{S}}^{\frac{1}{2}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}\, \widehat{\mathbf{S}}^{\frac{1}{2}} \tag{14}$$

where $\mathbf{P}_{\widehat{\mathbf{C}}_1} = \widehat{\mathbf{C}}_1(\widehat{\mathbf{C}}_1^H\widehat{\mathbf{C}}_1)^{-1}\widehat{\mathbf{C}}_1^H$ is the projection matrix onto the space of $\widehat{\mathbf{C}}_1$. Moreover, let $\mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp} = \mathbf{I} - \mathbf{P}_{\widehat{\mathbf{C}}_1}$ be the projection onto the space orthogonal to $\widehat{\mathbf{C}}_1$. We have

$$\mathrm{Tr}(\mathbf{Q}^*) = \mathrm{Tr}(\widehat{\mathbf{S}}^{\frac{1}{2}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}\, \widehat{\mathbf{S}}^{\frac{1}{2}})$$

$$= \mathrm{Tr}(\widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}) \tag{15}$$

$$= \mathrm{Tr}(\widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}\mathbf{P}_{\widehat{\mathbf{C}}_1}) \tag{16}$$

$$= \mathrm{Tr}(\mathbf{P}_{\widehat{\mathbf{C}}_1}\widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}) \tag{17}$$

where (15) comes from the fact that $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$, and (16) is because $\mathbf{P}_{\widehat{\mathbf{C}}_1} = \mathbf{P}_{\widehat{\mathbf{C}}_1}\mathbf{P}_{\widehat{\mathbf{C}}_1}$. Similarly we have

$$\mathrm{Tr}(\widehat{\mathbf{S}}) = \mathrm{Tr}\left((\mathbf{P}_{\widehat{\mathbf{C}}_1} + \mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp})\, \widehat{\mathbf{S}}\, (\mathbf{P}_{\widehat{\mathbf{C}}_1} + \mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp})\right)$$

$$= \mathrm{Tr}(\mathbf{P}_{\widehat{\mathbf{C}}_1}\, \widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}) + \mathrm{Tr}(\mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}\, \widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}) \tag{18}$$

$$= \mathrm{Tr}(\mathbf{Q}_{\widehat{S}}^*) + \mathrm{Tr}(\mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}\, \widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}) \tag{19}$$

where (19) results from (17).

Lemma 1. For the optimal $\widehat{\mathbf{S}}$, we have $\mathrm{span}\{\widehat{\mathbf{C}}_1\} = \mathrm{span}\{\widehat{\mathbf{S}}\}$.

Proof: : The proof is obtained using (19), and by noting that for the optimal $\widehat{\mathbf{S}}$ we must have $\mathrm{Tr}(\widehat{\mathbf{S}}) = \mathrm{Tr}(\mathbf{Q}^*) = P_t$. This means that we must have $\mathrm{Tr}(\mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}\, \widehat{\mathbf{S}}\, \mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}) = 0$, or equivalently $\mathbf{P}_{\widehat{\mathbf{C}}_1}^{\perp}\, \widehat{\mathbf{S}} = \underline{0}$, which completes the proof. ∎

Using Lemma 1 in (14) we have

$$\mathbf{Q}^* = \widehat{\mathbf{S}}. \tag{20}$$

The following lemma reveals another property of the optimal input covariance matrix under the average power constraint.

Lemma 2. For the optimal $\widehat{\mathbf{S}}$, i.e. $\mathbf{Q}^*$, the pencil $(\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}$, $\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I})$ has no generalized eigenvalue less than one:

$$\widehat{\mathbf{C}}^H\left[\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}\right]\widehat{\mathbf{C}} = \left[\begin{array}{cc} \widehat{\mathbf{\Lambda}}_1 & \underline{0} \\ \underline{0} & \mathbf{I} \end{array}\right] \tag{21}$$

$$\widehat{\mathbf{C}}^H\left[\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}\right]\widehat{\mathbf{C}} = \left[\begin{array}{cc} \mathbf{I} & \underline{0} \\ \underline{0} & \mathbf{I} \end{array}\right] \tag{22}$$

where $\widehat{\mathbf{\Lambda}}_2 = \mathbf{I}$ corresponds to the generalized eigenvalues equal to one.

Proof: We note that any vector which lies in the null space of $\widehat{\mathbf{S}}$ can be a generalized eigenvector of the pencil $(\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}$, $\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I})$, with a generalized eigenvalue equal to 1. Such vectors span the null space of $\widehat{\mathbf{S}}$, i.e., $\mathrm{span}\{\widehat{\mathbf{C}}_2\} = \mathrm{span}\{\widehat{\mathbf{S}}\}^\perp$. On the other hand, from Lemma 1, $\mathrm{span}\{\widehat{\mathbf{C}}_1\} = \mathrm{span}\{\widehat{\mathbf{S}}\}$. Thus for the optimal $\widehat{\mathbf{S}}$, all generalized eigenvectors of the pencil $(\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}$, $\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I})$ correspond to generalized eigenvalues either bigger than or equal to 1. ∎

Let b denote number of generalized eigenvalues of the pencil $(\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I}$, $\widehat{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\widehat{\mathbf{S}}^{\frac{1}{2}}+\mathbf{I})$ that are strictly bigger than 1, where again $\widehat{\mathbf{S}} = \mathbf{Q}^*$ represents the optimal input covariance matrix that attains the secrecy capacity under the average power constraint given by (11). From Lemma 1, we have $\mathrm{rank}(\mathbf{Q}^*) = \mathrm{rank}(\widehat{\mathbf{C}}_1) = b$.

Theorem 1. For the optimal $\mathbf{Q}^*$ we have

$$\mathrm{rank}(\mathbf{Q}^*) \leq m \tag{23}$$

where m is the number of positive eigenvalues of the matrix $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$.

Proof: Subtracting (22) from (21), a straightforward computation yields

$$\widehat{\mathbf{S}}^{\frac{1}{2}}\left[\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}\right]\widehat{\mathbf{S}}^{\frac{1}{2}} = \widehat{\mathbf{C}}^{-H}\left[\begin{array}{cc} \widehat{\mathbf{\Lambda}}_1 - \mathbf{I} & \underline{0} \\ \underline{0} & \underline{0} \end{array}\right]\widehat{\mathbf{C}}^{-1} \succeq \underline{0}. \tag{24}$$

From (24), we note that $\widehat{\mathbf{S}}^{\frac{1}{2}}\left[\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}\right]\widehat{\mathbf{S}}^{\frac{1}{2}} \succeq \underline{0}$, from which it follows that $\mathrm{rank}(\mathbf{Q}^*) \leq m$. ∎

Remark 2. From Theorem 1, one can easily confirm that the optimal $\mathbf{Q}^*$ can be full rank only in the case that $m = n_t$, i.e. $\mathbf{H}^H\mathbf{H} \succ \mathbf{G}^H\mathbf{G}$. For all other scenarios, the optimal $\mathbf{Q}^*$ will be low rank. The authors

of [3], [5] use the Karush-Kuhn-Tucker (KKT) conditions on problem (2) to make a similar statement, but they do not show what the rank of the optimal $\mathbf{Q}^*$ will be.

The following lemma will be used for the computations in the next section.

Lemma 3. For the case of $\mathbf{H}^H\mathbf{H} \succ \mathbf{G}^H\mathbf{G}$, for any $n_t \times n_t$ matrix $\mathbf{S} \succeq \underline{0}$, all the generalized eigenvalues of the pencil $(\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}} + \mathbf{I}$, $\mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\mathbf{S}^{\frac{1}{2}} + \mathbf{I})$ are strictly bigger than 1, i.e. $\mathbf{\Lambda} \succ \mathbf{I}$, *iff* $\mathbf{S}$ is full rank, i.e. $\mathbf{S} \succ \underline{0}$.

Proof: The claim is easily proved by considering the rank of both sides of (24). ∎

IV. CHARACTERIZATION OF THE OPTIMAL FULL-RANK SOLUTION

In this section, we characterize the secrecy capacity under the average power constraint for a particular class of MIMO Gaussian wiretap channel where the optimal solution $\mathbf{Q}^*$ is full rank. While necessary conditions for a full-rank $\mathbf{Q}^*$ were characterized in the previous section, here we derive sufficient conditions as well.

We begin by rewriting problem (2) here:

$$\mathcal{C}_{sec}(P_t) = \max_{\mathbf{Q} \succeq \underline{0},\ \mathrm{Tr}(\mathbf{Q}) = P_t} \log|\mathbf{HQH}^H + \mathbf{I}| - \log|\mathbf{GQG}^H + \mathbf{I}| . \tag{25}$$

The Lagrangian associated with this problem is given by

$$\mathcal{L} = \log|\mathbf{HQH}^H + \mathbf{I}| - \log|\mathbf{GQG}^H + \mathbf{I}| - \mu(\mathrm{Tr}(\mathbf{Q}) - P_t) + \mathrm{Tr}(\mathbf{MQ}) \tag{26}$$

where $\mu > 0$ and $\mathbf{M} \succeq \underline{0}$ are the Lagrange multipliers. The optimal $\mathbf{Q}^*$ must satisfy the following KKT conditions:

$$\mathbf{H}^H(\mathbf{HQ}^*\mathbf{H}^H + \mathbf{I})^{-1}\mathbf{H} - \mathbf{G}^H(\mathbf{GQ}^*\mathbf{G}^H + \mathbf{I})^{-1}\mathbf{G} = \mu\mathbf{I} - \mathbf{M} \tag{27}$$

$$\mu(\mathrm{Tr}(\mathbf{Q}^*) - P_t) = 0 \tag{28}$$

$$\mathbf{Q}^*\mathbf{M} = \mathbf{MQ}^* = \underline{0} . \tag{29}$$

Using the matrix inversion lemma [14], (27) can be written as

$$(\mathbf{H}^H\mathbf{HQ}^* + \mathbf{I})^{-1}\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}(\mathbf{Q}^*\mathbf{G}^H\mathbf{G} + \mathbf{I})^{-1} = \mu\mathbf{I} - \mathbf{M} . \tag{30}$$

Left multiplication by $(\mathbf{H}^H\mathbf{HQ}^* + \mathbf{I})$ and right multiplication by $(\mathbf{Q}^*\mathbf{G}^H\mathbf{G} + \mathbf{I})$ of both sides of (30) yields

$$\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} = \mu\left(\mathbf{H}^H\mathbf{HQ}^* + \mathbf{I}\right)\left(\mathbf{Q}^*\mathbf{G}^H\mathbf{G} + \mathbf{I}\right) - \mathbf{M} \tag{31}$$

$$= \mu\left(\mathbf{G}^H\mathbf{GQ}^* + \mathbf{I}\right)\left(\mathbf{Q}^*\mathbf{H}^H\mathbf{H} + \mathbf{I}\right) - \mathbf{M} , \tag{32}$$

where in obtaining (31) we have used the KKT condition (29), and Eq. (32) comes from the fact that (31) is Hermitian.

We are considering problem (25) for the case that $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$ is strictly positive definite, i.e. $\mathbf{H}^H\mathbf{H} \succ \mathbf{G}^H\mathbf{G}$, since this is the necessary condition for having a full rank optimal $\mathbf{Q}^*$. As we characterize the full rank $\mathbf{Q}^*$, the sufficient conditions are revealed as well.

Remark 3. By following exactly the same steps as in the proof of [3, Proposition 5], one can easily show that for the case of $\mathbf{H}^H\mathbf{H} \succ \mathbf{G}^H\mathbf{G}$, the optimization problem (25) is convex[2] in $\mathbf{Q}$.

Thus for the case of interest, the KKT conditions (28), (29) and (31) are necessary and sufficient conditions for the optimality of $\mathbf{Q}^*$. In other words, any $\mathbf{Q} \succeq \underline{0}$ that satisfies those conditions is an optimal solution for the problem (25). By the KKT condition (29), a full rank $\mathbf{Q}^*$ follows that $\mathbf{M} = \underline{0}$. Thus, (31) and (32) simplify to

$$\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} = \mu\,(\mathbf{H}^H\mathbf{H}\mathbf{Q}^* + \mathbf{I})\,(\mathbf{Q}^*\mathbf{G}^H\mathbf{G} + \mathbf{I}) \tag{33}$$

$$= \mu\,(\mathbf{G}^H\mathbf{G}\mathbf{Q}^* + \mathbf{I})\,(\mathbf{Q}^*\mathbf{H}^H\mathbf{H} + \mathbf{I}) . \tag{34}$$

Lemma 4. Let the diagonal matrix $\mathbf{D}$ and the unitary matrix $\boldsymbol{\Phi}_{\bar{\mathbf{s}}}$ respectively denote the eigenvalue and eigenvector matrices of $(\bar{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I})$, where we set $\bar{\mathbf{S}} = (\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G})^{-1}$:

$$(\bar{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I}) = \boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,\mathbf{D}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H. \tag{35}$$

Then we have

$$\mathbf{H}^H\mathbf{H} = \bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,\mathbf{D}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}} \tag{36}$$

$$\mathbf{G}^H\mathbf{G} = \bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{D} - \mathbf{I})\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}} . \tag{37}$$

Proof: Eq. (37) comes directly from (35). Please refer to Appendix A for details on obtaining (36). ∎

Using (36) and (37) in (33), after some simplification we have

$$\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}} = \mu\,(\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,\mathbf{D}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}}\,\mathbf{Q}^* + \mathbf{I})\,(\mathbf{Q}^*\,\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{D} - \mathbf{I})\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}} + \mathbf{I})$$

$$= \mu\,\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{D}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}}\,\mathbf{Q}^* + \boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{\frac{1}{2}})\,(\mathbf{Q}^*\,\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{D} - \mathbf{I}) + \bar{\mathbf{S}}^{\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}})\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}}$$

$$= \mu\,\bar{\mathbf{S}}^{-\frac{1}{2}}\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{D}\,\mathbf{W} + \mathbf{I})\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\,(\mathbf{W}(\mathbf{D} - \mathbf{I}) + \mathbf{I})\,\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{-\frac{1}{2}} , \tag{38}$$

[2]In fact, the optimization problem (25) is convex in $\mathbf{Q}$ when $\mathbf{H}^H\mathbf{H} \succeq \mathbf{G}^H\mathbf{G}$.

where in (38) we defined

$$\mathbf{W} = \mathbf{\Phi}_{\bar{\mathbf{s}}}^{H} \bar{\mathbf{S}}^{-\frac{1}{2}} \, \mathbf{Q}^{*} \bar{\mathbf{S}}^{-\frac{1}{2}} \mathbf{\Phi}_{\bar{\mathbf{s}}} \, . \tag{39}$$

From (38), we get

$$\mathbf{I} = \mu \, (\mathbf{D} \, \mathbf{W} + \mathbf{I}) \, \mathbf{\Phi}_{\bar{\mathbf{s}}}^{H} \bar{\mathbf{S}} \mathbf{\Phi}_{\bar{\mathbf{s}}} \, (\mathbf{W}(\mathbf{D} - \mathbf{I}) + \mathbf{I}) \, . \tag{40}$$

Let $\mathbf{X} \succ \underline{0}$ and $\mathbf{Y} \succ \underline{0}$ be two diagonal matrices, which will be defined soon. Left multiplication by $\mathbf{X}$ and right multiplication by $\mathbf{Y}$ of both sides of (40) gives

$$\mathbf{XY} = \mu \, (\mathbf{XD} \, \mathbf{W} + \mathbf{X}) \, \mathbf{\Phi}_{\bar{\mathbf{s}}}^{H} \bar{\mathbf{S}} \mathbf{\Phi}_{\bar{\mathbf{s}}} \, (\mathbf{W}(\mathbf{D} - \mathbf{I})\mathbf{Y} + \mathbf{Y}) \, . \tag{41}$$

We find $\mathbf{X}$ and $\mathbf{Y}$ by solving the following set of equations

$$\begin{aligned} \mathbf{XY} &= \mathbf{I} \\ \mathbf{XD} &= (\mathbf{D} - \mathbf{I})\mathbf{Y} \end{aligned} \tag{42}$$

which results in

$$\mathbf{X} = \mathbf{Y}^{-1} = \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{\frac{1}{2}} \, . \tag{43}$$

Substituting (43) into (41), we have

$$\begin{aligned} \mathbf{I} &= \mu \, (\mathbf{XD} \, \mathbf{W} + \mathbf{X}) \, \mathbf{\Phi}_{\bar{\mathbf{s}}}^{H} \bar{\mathbf{S}} \mathbf{\Phi}_{\bar{\mathbf{s}}} \, \left(\mathbf{W} \, \mathbf{DX} + \mathbf{X}^{-1}\right) \\ &= \mu \, \left(\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + \mathbf{X}^{2}\mathbf{D}\right) \, \mathbf{X}^{-1}\mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^{H}\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\mathbf{X}^{-1} \, (\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + \mathbf{D}) \\ &= \mu \, (\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + (\mathbf{D} - \mathbf{I})) \, \mathbf{X}^{-1}\mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^{H}\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\mathbf{X}^{-1} \, (\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + \mathbf{D}) \, . \end{aligned} \tag{44}$$

where we have used (43) in obtaining (44).

Using an approach similar to what we used to obtain (44) from (33), one can show that[3]

$$\mathbf{I} = \mu \, (\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + \mathbf{D}) \, \mathbf{X}^{-1}\mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^{H}\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\mathbf{X}^{-1} \, (\mathbf{XD} \, \mathbf{W} \, \mathbf{DX} + (\mathbf{D} - \mathbf{I})) \, . \tag{45}$$

Define $\mathbf{K}$ as

$$\mathbf{K} = \mathbf{XD} \, \mathbf{W} \, \mathbf{DX} = \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{\frac{1}{2}} \mathbf{D} \, \mathbf{W} \, \mathbf{D} \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{\frac{1}{2}} \, , \tag{46}$$

where $\mathbf{D}$ and $\mathbf{W}$ are respectively given by (35) and (39). Moreover, let

$$\bar{\mathbf{\Sigma}} = \mathbf{X}^{-1}\mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^{H}\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\mathbf{X}^{-1} = \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}} \mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^{H}\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1} \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}} \, . \tag{47}$$

---

[3]Clearly (45) is also trivially obtained from (44).

Using (46) and (47) in (44) and (45), we have

$$
\begin{aligned}
\mathbf{I} &= \mu \, (\mathbf{K} + \mathbf{D}) \, \bar{\boldsymbol{\Sigma}} \, (\mathbf{K} + (\mathbf{D} - \mathbf{I})) \\
&= \mu \, (\mathbf{K} + (\mathbf{D} - \mathbf{I})) \, \bar{\boldsymbol{\Sigma}} \, (\mathbf{K} + \mathbf{D}) \; .
\end{aligned}
\tag{48}
$$

We note from (48) that

$$
\mu \, \bar{\boldsymbol{\Sigma}} = (\mathbf{K} + \mathbf{D})^{-1} \, (\mathbf{K} + (\mathbf{D} - \mathbf{I}))^{-1} = (\mathbf{K} + (\mathbf{D} - \mathbf{I}))^{-1} \, (\mathbf{K} + \mathbf{D})^{-1} \; .
\tag{49}
$$

This result implies that, for the optimal $\mathbf{Q}^*$, $(\mathbf{K} + (\mathbf{D} - \mathbf{I}))^{-1}$, $(\mathbf{K} + \mathbf{D})^{-1}$ and $\bar{\boldsymbol{\Sigma}}$ commute and have the same eigenvectors [14].

Denote the eigenvalue decomposition of $\bar{\boldsymbol{\Sigma}}$ as

$$
\bar{\boldsymbol{\Sigma}} = \mathbf{U} \, \boldsymbol{\Omega} \, \mathbf{U}^H \; .
\tag{50}
$$

Based on the argument made after (49), we have

$$
\begin{aligned}
\mathbf{K} + (\mathbf{D} - \mathbf{I}) &= \mathbf{U} \, \boldsymbol{\Lambda}_1 \, \mathbf{U}^H \\
\mathbf{K} + \mathbf{D} &= \mathbf{U} \, \boldsymbol{\Lambda}_2 \, \mathbf{U}^H \; ,
\end{aligned}
\tag{51}
$$

where one can easily confirm that $\boldsymbol{\Lambda}_2 = \boldsymbol{\Lambda}_1 + \mathbf{I}$. By replacing (50) and (51) in (48), and noting that $\mathbf{U}^H \mathbf{U} = \mathbf{U} \mathbf{U}^H = \mathbf{I}$, and that $\boldsymbol{\Lambda}_1$, $\boldsymbol{\Lambda}_2$ and $\boldsymbol{\Omega}$ are all diagonal, we have

$$
\mathbf{I} = \mu \, \mathbf{U} \boldsymbol{\Lambda}_1 \, \boldsymbol{\Omega} \, \boldsymbol{\Lambda}_2 \mathbf{U}^H = \mu \, \boldsymbol{\Lambda}_1 \, \boldsymbol{\Omega} \, \boldsymbol{\Lambda}_2 = \mu \, (\boldsymbol{\Lambda}_1^2 + \boldsymbol{\Lambda}_1) \, \boldsymbol{\Omega} \; .
\tag{52}
$$

Recall that the unknown parameters in (52) are the diagonal matrix $\boldsymbol{\Lambda}_1$ and the scalar $\mu > 0$. Let λ_{i1} and ω_i denote the ith diagonal element of $\boldsymbol{\Lambda}_1$ and $\boldsymbol{\Omega}$, respectively. From (52), we can solve for λ_{i1} and obtain

$$
\lambda_{i1} = \frac{1}{2} \left(-1 + \sqrt{1 + \frac{4}{\mu \omega_i}} \right) \; ,
\tag{53}
$$

where the Lagrange multiplier $\mu > 0$ is chosen to satisfy the power constraint $\mathrm{Tr}(\mathbf{Q}^*) = P_t$, as will be explained below.

Theorem 2. The optimal full-rank input covariance matrix that attains the secrecy capacity for the average power constraint is given by

$$
\mathbf{Q}^* = \bar{\mathbf{S}}^{\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}} \left(\mathbf{I} - \mathbf{D}^{-1} \right)^{-\frac{1}{2}} \mathbf{D}^{-1} \left(\mathbf{U} \boldsymbol{\Lambda}_1 \mathbf{U}^H + \mathbf{I} - \mathbf{D} \right) \mathbf{D}^{-1} \left(\mathbf{I} - \mathbf{D}^{-1} \right)^{-\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H \bar{\mathbf{S}}^{\frac{1}{2}}
\tag{54}
$$

iff

- $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succ \underline{0}$, and

- for the given P_t,

$$\mathbf{U}\boldsymbol{\Lambda}_1\mathbf{U}^H \succ \mathbf{D} - \mathbf{I} \tag{55}$$

where $\boldsymbol{\Lambda}_1$ is given by (53).

Proof: The proof is obtained by obtaining $\mathbf{K}$ from (51), and substituting it back into (46) and (39) via straightforward computations. ∎

To fully characterize $\mathbf{Q}^*$, one must obtain the Lagrange multiplier $\mu > 0$ such that $\mathrm{Tr}(\mathbf{Q}^*) = P_t$. As (53) shows, $\mathrm{Tr}(\mathbf{Q}^*)$ is monotonically decreasing with μ:

$$\lim_{\mu \to 0} \mathrm{Tr}(\mathbf{Q}^*) = \infty, \qquad \text{and} \qquad \lim_{\mu \to \infty} \mathrm{Tr}(\mathbf{Q}^*) = -\mathrm{Tr}(\bar{\mathbf{S}}^{\frac{1}{2}}\boldsymbol{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{\frac{1}{2}}) < 0 \, .$$

Thus for any transmit power P_t, there exists a Lagrange multiplier $\mu > 0$ for which $\mathrm{Tr}(\mathbf{Q}^*) = P_t$. The appropriate value of μ can be easily found using, for example, the bisection method.

Note that Theorem 2 also reveals the necessary and sufficient conditions for having a full rank optimal $\mathbf{Q}^*$. While for any transmit power P_t one can find a Lagrange multiplier $\mu > 0$ for which $\mathrm{Tr}(\mathbf{Q}^*) = P_t$ and $\mathbf{U}\boldsymbol{\Lambda}_1\mathbf{U}^H \succ \underline{0}$, to have a full rank $\mathbf{Q}^*$, (55) must be satisfied. Also recall from (74) that $\mathbf{D} - \mathbf{I} \succeq \underline{0}$. The flowchart in Fig. 1 summarizes the steps required to calculate the optimal full-rank $\mathbf{Q}^*$ for the case of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} \succ \underline{0}$.

Remark 4. By replacing $\mathbf{Q}^*$ given by (54) into (25), the optimal input covariance matrix $\mathbf{Q}^*$ given by Theorem 2 attains the secrecy capacity

$$\mathcal{C}_{sec}(P_t) = \log|\boldsymbol{\Lambda}_1| - \log|\boldsymbol{\Lambda}_1 + \mathbf{I}| + \log|\mathbf{D}| - \log|\mathbf{D} - \mathbf{I}|$$

$$= \log|\mathbf{I} - (\boldsymbol{\Lambda}_1 + \mathbf{I})^{-1}| + \log|\mathbf{I} + (\mathbf{D} - \mathbf{I})^{-1}| \, , \tag{56}$$

where $\boldsymbol{\Lambda}_1$ and $\mathbf{D}$ are diagonal matrices, respectively given by (53) and (35). Note that while both $\log$ terms in (56) return non-negative values, the first term depends on both the channels and the power P_t, while the second term just depends on the channels (see Lemma 5).

In the following we show that for the case of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} \succeq \underline{0}$, i.e. when at least one of the eigenvalues of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$ is zero, there exists an equivalent wiretap channel with the same secrecy capacity of the original wiretap channel. For this case, the secrecy capacity is characterized too.

Theorem 3. For the MIMO Gaussian wiretap channel defined by direct and cross channels $\mathbf{H}$ and $\mathbf{G}$ respectively, with $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} \succeq \underline{0}$, there exists an equivalent wiretap channel with $\mathbf{H}_{eq}$ and $\mathbf{G}_{eq}$ such that $\mathbf{H}_{eq}^H\mathbf{H}_{eq} - \mathbf{G}_{eq}^H\mathbf{G}_{eq} \succ \underline{0}$, and $\mathcal{C}_{sec}(\mathbf{H}, \mathbf{G}, P_t) = \mathcal{C}_{sec}(\mathbf{H}_{eq}, \mathbf{G}_{eq}, P_t)$.

$$\boxed{\mathbf{H}, \mathbf{G}, P_t}$$

Step 1: Create $\bar{\mathbf{S}}$

$$\boxed{\bar{\mathbf{S}} = (\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G})^{-1}}$$

Step 2: Calculate $\mathbf{\Phi}_{\bar{\mathbf{s}}}$ and $\mathbf{D}$ from (35)

$$\boxed{[\mathbf{\Phi}_{\bar{\mathbf{s}}}, \mathbf{D}] = \mathrm{eig}(\bar{\mathbf{S}}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I})}$$

Step 3: Calculate $\bar{\mathbf{\Sigma}}$ from (47)

$$\boxed{\bar{\mathbf{\Sigma}} = \left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}} \mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}}}$$

Step 4: Calculate $\mathbf{U}$ and $\mathbf{\Omega}$ from (50)

$$\boxed{[\mathbf{U}, \mathbf{\Omega}] = \mathrm{eig}(\bar{\mathbf{\Sigma}})}$$

Step 5: Derive $\mathbf{\Lambda}_1$ from (53)

$$\boxed{\mathbf{\Lambda}_1 = \frac{1}{2}\left(-\mathbf{I} + (\mathbf{I} + \frac{4}{\mu}\mathbf{\Omega}^{-1})^{\frac{1}{2}}\right)}$$

Step 6: Calculate $\mathbf{Q}^*$ according to (54)

$$\boxed{\mathbf{Q}^* = \bar{\mathbf{S}}^{\frac{1}{2}}\mathbf{\Phi}_{\bar{\mathbf{s}}}\mathbf{D}^{-1}\left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}}\left(\mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H + \mathbf{I} - \mathbf{D}\right)\left(\mathbf{I} - \mathbf{D}^{-1}\right)^{-\frac{1}{2}}\mathbf{D}^{-1}\mathbf{\Phi}_{\bar{\mathbf{s}}}^H\bar{\mathbf{S}}^{\frac{1}{2}}}$$

Step 7: Find μ

$$\boxed{\mu|_{\mathrm{Tr}(\mathbf{Q}^*)=P_t}}$$

Step 8: Check the validity of $\mathbf{Q}^*$

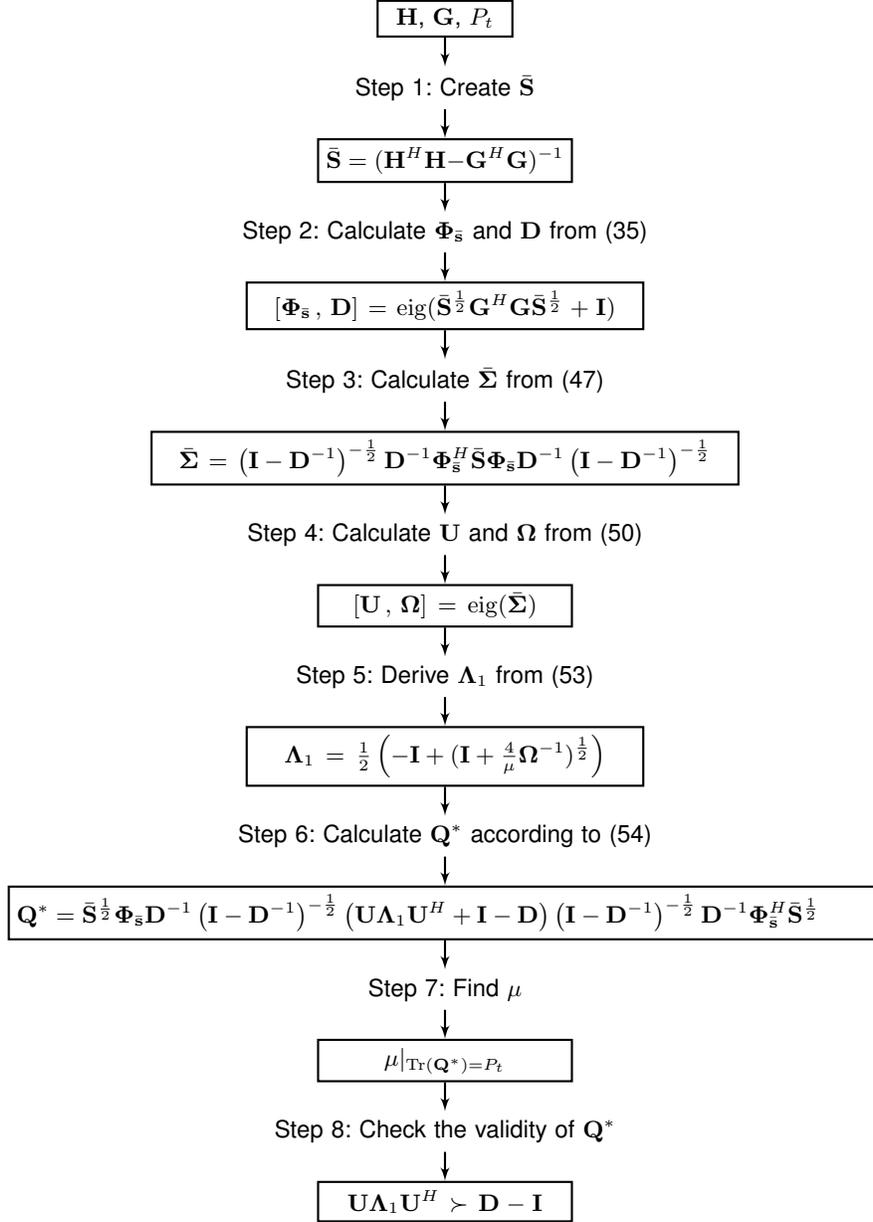$$\boxed{\mathbf{U}\mathbf{\Lambda}_1\mathbf{U}^H \succ \mathbf{D} - \mathbf{I}}$$

Fig. 1: Flowchart for obtaining the optimal full-rank $\mathbf{Q}^*$.

Proof: Please refer to Appendix B for the proof and the characterization of the equivalent channels $\mathbf{H}_{eq}$ and $\mathbf{G}_{eq}$. ∎

Although at this time a proof is unavailable, we conjecture that the secrecy capacity of any general MIMO wiretap channel is given by an equivalent wiretap channel with $\mathbf{H}_{eq}$ and $\mathbf{G}_{eq}$ such that $\mathbf{H}_{eq}^H\mathbf{H}_{eq}^H -$

$\mathbf{G}_{eq}^H \mathbf{G}_{eq} \succ \underline{0}$, and $\mathcal{C}_{sec}(\mathbf{H}, \mathbf{G}, P_t) = \mathcal{C}_{sec}(\mathbf{H}_{eq}, \mathbf{G}_{eq}, P_t)$. If one cannot find such an equivalent wiretap channel, the secrecy capacity of the main channel is zero, i.e., $\mathcal{C}_{sec}(\mathbf{H}, \mathbf{G}, P_t) = 0$. Besides the case of $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succeq \underline{0}$, the case for which the optimal input covariance matrix is rank 1 is another case supporting this conjecture. For the rank 1 case, it is shown in [6] that the optimal input covariance matrix is given by $\mathbf{Q}^* = P_t \, \mathbf{u} \mathbf{u}^H$, where $\mathbf{u}$ is the normalized principal eigenvector corresponding to the largest eigenvalue λ_1 of the pencil $(\mathbf{I} + P_t \mathbf{H}^H \mathbf{H} \,, \, \mathbf{I} + P_t \mathbf{G}^H \mathbf{G})$. For this case, if $\lambda_1 > 1$, the secrecy capacity is $\mathcal{C}_{sec}(\mathbf{H}, \mathbf{G}, P_t) = \log(\lambda_1)$, and the equivalent wiretap channel is defined such that $\mathbf{h}_{eq}^H \mathbf{h}_{eq} = \mathbf{u}^H \mathbf{H}^H \mathbf{H} \mathbf{u}$ and $\mathbf{g}_{eq}^H \mathbf{g}_{eq} = \mathbf{u}^H \mathbf{G}^H \mathbf{G} \mathbf{u}$.

V. REMARKS REGARDING $\mathbf{Q}^*$

This section discusses some interesting points regarding the optimal solution in (54). For the following observations, one can assume when required that both conditions for a full-rank $\mathbf{Q}^*$ given in Theorem (2) are satisfied. Let γ_i, $i = 1, \cdots, n_t$, be the generalized eigenvalues of the pencil $(\mathbf{H}^H \mathbf{H} \,, \, \mathbf{G}^H \mathbf{G})$. Then from the definition [15], $\gamma_i = \sigma_i^2$, where σ_i is the ith generalized singular value of $(\mathbf{H} \,, \, \mathbf{G})$.

Lemma 5. The second term in the secrecy capacity expression (56) is only channel dependent and is equal to $\sum_{i=1}^{n_t} \log(\sigma_i^2)$.

Proof: From (72) we have:

$$
\begin{aligned}
\mathbf{H}^H \mathbf{H} \, \bar{\mathbf{S}}^{\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}} &= \bar{\mathbf{S}}^{-\frac{1}{2}} \, \boldsymbol{\Phi}_{\bar{\mathbf{s}}} \, \mathbf{D} \\
&= \mathbf{G}^H \mathbf{G} \, \bar{\mathbf{S}}^{\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}} \, (\mathbf{D} - \mathbf{I})^{-1} \, \mathbf{D} \\
&= \mathbf{G}^H \mathbf{G} \, \bar{\mathbf{S}}^{\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}} \left(\mathbf{I} + (\mathbf{D} - \mathbf{I})^{-1} \right) \,,
\end{aligned}
\tag{57}
$$

where (57) comes from (73). Thus, from the definition [14], the generalized eigenvalue matrix of $(\mathbf{H}^H \mathbf{H} \,, \, \mathbf{G}^H \mathbf{G})$ is $\left(\mathbf{I} + (\mathbf{D} - \mathbf{I})^{-1} \right)$, which completes the proof. ∎

Note that the definition of singular values here is slightly different from what is given in [4]. Here σ_i may be ∞, while this is not the case in [4]. More precisely, from (73) for the case of $\text{rank}(\mathbf{G}) = n_e < n_t$, $n_t - n_e$ diagonal elements of $\mathbf{D}$ are equal to one, as mentioned in Remark 5. The σ_i corresponding to $d_i = 1$ tends to ∞.

Lemma 6. In the high SNR scenario $(P_t \to \infty)$ and for the case of $\text{rank}(\mathbf{G}) = n_t$, the asymptotic form of the exact secrecy capacity (56) is simply given by

$$
\mathcal{C}_{sec} = \log |\mathbf{I} + (\mathbf{D} - \mathbf{I})^{-1}| = \sum_{i=1}^{n_t} \log(\sigma_i^2) \,.
\tag{58}
$$

Proof: For $P_t \to \infty$, the Lagrange parameter satisfies $\mu \to 0$, as mentioned after Theorem 2. Moreover, for the case of $\mathrm{rank}(\mathbf{G}) = n_t$, the matrix $\bar{\boldsymbol{\Sigma}}$ given by (47) will have finite-valued eigenvalues. Thus as $\mu \to 0$, the elements of the diagonal matrix $\boldsymbol{\Lambda}_1$, given by (53), go to ∞ ($\lambda_{i1} \to \infty$). Consequently, the first $\log$ term in (56) disappears as $P_t \to \infty$. ∎

It is also interesting to consider the optimal solution in (54) for the case that the eavesdropper's channel is very weak, e.g. $\mathbf{G} = \underline{0}$. For this specific case, the wiretap channel simplifies to a point-to-point MIMO Gaussian link, where the optimal input covariance matrix under the average power constraint is known to be $\boldsymbol{\Phi}_H \left(\frac{1}{\mu} \mathbf{I} - \boldsymbol{\Lambda}_H^{-1} \right)^+ \boldsymbol{\Phi}_H^H$, and is found via the standard water-filling solution, where unitary $\boldsymbol{\Phi}_H$ and diagonal $\boldsymbol{\Lambda}_H$ are obtained from the eigenvalue decomposition $\mathbf{H}^H \mathbf{H} = \boldsymbol{\Phi}_H \boldsymbol{\Lambda}_H \boldsymbol{\Phi}_H^H$.

Lemma 7. For the case of $\mathbf{G} = \underline{0}$, the optimal solution in (54) simplifies to the conventional water-filling solution, where

$$\mathbf{Q}^* = \boldsymbol{\Phi}_H \left(\mu^{-1} \mathbf{I} - \boldsymbol{\Lambda}_H^{-1} \right)^+ \boldsymbol{\Phi}_H^H . \tag{59}$$

Proof: Using (72) and via simple calculations, we note that for any $\mathbf{G}$, Eq. (54) can be rewritten as

$$\mathbf{Q}^* = \bar{\mathbf{S}}^{\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}} \left(\mathbf{I} - \mathbf{D}^{-1} \right)^{-\frac{1}{2}} \mathbf{D}^{-1} \mathbf{U} \boldsymbol{\Lambda}_1 \mathbf{U}^H \mathbf{D}^{-1} \left(\mathbf{I} - \mathbf{D}^{-1} \right)^{-\frac{1}{2}} \boldsymbol{\Phi}_{\bar{\mathbf{s}}}^H \bar{\mathbf{S}}^{\frac{1}{2}} - (\mathbf{H}^H \mathbf{H})^{-1} . \tag{60}$$

From (73), when $\mathbf{G} \to \underline{0}$ then $\mathbf{D} \to \mathbf{I}$. Next from (72), $\boldsymbol{\Phi}_{\bar{\mathbf{s}}} \to \boldsymbol{\Phi}_H$. Using these facts in (47), and after some straightforward calculations, we have $\bar{\boldsymbol{\Sigma}} \to (\mathbf{D} - \mathbf{I})^{-1} \mathbf{D}^{-1} \boldsymbol{\Lambda}_H^{-1} = \boldsymbol{\Omega}$, and $\mathbf{U} \to \mathbf{I}$ in (50). Using these in (60), we have

$$\begin{aligned}
\mathbf{Q}^* &\to \boldsymbol{\Phi}_H (\mathbf{D} - \mathbf{I})^{-1} \mathbf{D}^{-1} \boldsymbol{\Lambda}_H^{-1} \boldsymbol{\Lambda}_1 \boldsymbol{\Phi}_H^H - (\mathbf{H}^H \mathbf{H})^{-1} \\
&\to \frac{1}{2} \boldsymbol{\Phi}_H (\mathbf{D} - \mathbf{I})^{-1} \boldsymbol{\Lambda}_H^{-1} \left(-\mathbf{I} + (\mathbf{I} + \frac{4}{\mu} (\mathbf{D} - \mathbf{I}) \boldsymbol{\Lambda}_H)^{\frac{1}{2}} \right) \boldsymbol{\Phi}_H^H - (\mathbf{H}^H \mathbf{H})^{-1} \\
&\to \frac{1}{\mu} \mathbf{I} - (\mathbf{H}^H \mathbf{H})^{-1} = \boldsymbol{\Phi}_H \left(\mu^{-1} \mathbf{I} - \boldsymbol{\Lambda}_H^{-1} \right)^+ \boldsymbol{\Phi}_H^H ,
\end{aligned} \tag{61}$$

where in obtaining (61) we used the fact that $\mathbf{D} \to \mathbf{I}$ when $\mathbf{G} \to \underline{0}$. ∎

VI. NUMERICAL RESULTS

In the first example, we consider a MIMO wiretap channel with $n_t = n_e = 2$, $n_t = 3$ and channel matrices given by

$$\mathbf{H} = \begin{bmatrix} 0.32 - 0.52i & 0.83 + 1.15i \\ 0.51 - 0.26i & 0.06 - 0.15i \\ -0.11 + 0.81i & 0.29 + 0.68i \end{bmatrix} , \tag{62}$$
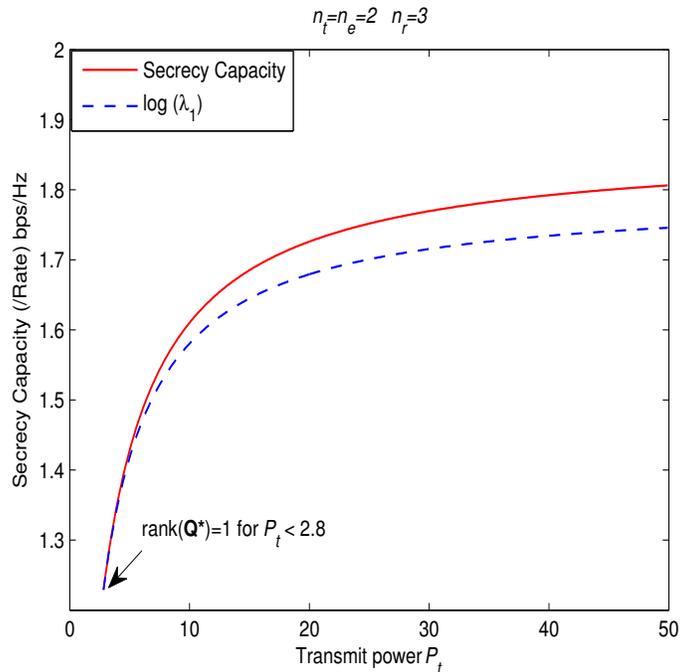
Fig. 2: Secrecy Capacity versus P_t for $n_t = n_e = 2$ and $n_r = 3$. Solid curve represents secrecy capacity and dotted curve indicates the achievable secrecy rate using a rank-one input covariance matrix.

$$\mathbf{G} = \begin{bmatrix} 0.03 - 0.70i & -0.32 - 0.32i \\ 0.24 - 0.11i & 1.36 + 0.18i \end{bmatrix}, \tag{63}$$

which satisfy $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succ \underline{0}$. Fig. 2 shows the secrecy capacity as a function of transmit power P_t. For comparison, the figure also depicts the achievable secrecy rate using the input covariance matrix $\mathbf{Q} = P_t \mathbf{u}\mathbf{u}^H$, which results to $R_{sec} = \log \lambda_1$, as shown in [4]-[6]. Note that in this example, the optimal $\mathbf{Q}^*$ is not full-rank for $P_t < 2.8$.

In Fig. 3 we consider another example of the case of $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succ \underline{0}$, here with $n_t = n_e = 3$, $n_t = 4$ and channel matrices given by

$$\mathbf{H} = \begin{bmatrix} 0.89 + 0.54i & -0.06 + 0.60i & 0.48 - 1.11i \\ 0.46 & -0.44 + 0.80i & -1.07 + 0.63i \\ 1.40 - 0.13i & 0.17 - 0.82i & 0.59 - 0.31i \\ 0.43 - 0.23i & 0.03 + 1.35i & 0.44 - 0.07i \end{bmatrix}, \tag{64}$$
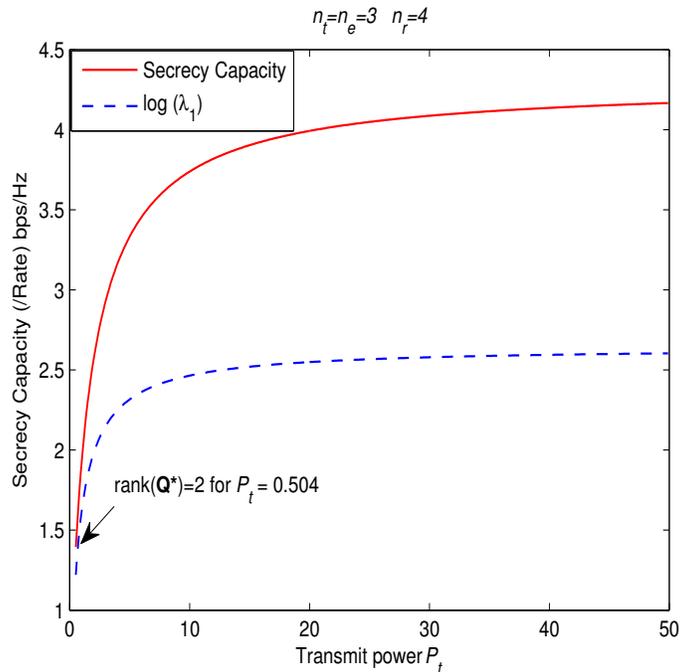
Fig. 3: Secrecy Capacity versus P_t for $n_t = n_e = 3$ and $n_r = 4$. Solid curve represents secrecy capacity and dotted curve indicates the achievable secrecy rate using a rank-one input covariance matrix.

$$\mathbf{G} = \begin{bmatrix} 0.46 - 0.59i & 0.24 - 0.01i & -0.37 + 0.15i \\ 0.51 - 0.63i & 0.58 + 0.51i & 0.86 - 0.47i \\ 0.17 - 0.24i & -0.83 + 0.51i & 0.04 - 0.64i \end{bmatrix}. \tag{65}$$

For this example, the optimal $\mathbf{Q}^*$ is only full-rank for $P_t > 0.5$.

Finally in Fig. 4, we compare the standard point-to-point capacity without secrecy constraints with the secrecy capacity given by (56). In this example, $P_t = 20$, direct channel $\mathbf{H}$ is given by (64) but the cross channel $\mathbf{G}$ is assumed to satisfy $\mathbf{G}^H \mathbf{G} = \alpha \mathbf{I}$, where α changes from 0 to 1.95 (note that $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succ \underline{0}$ only for $\alpha \leq 1.95$). As predicted, the secrecy capacity achieved by the derived $\mathbf{Q}^*$ in (54) approaches the standard capacity as $\mathbf{G} \to \underline{0}$. It is interesting to note that even for very small values of α, the difference between the standard capacity and secrecy capacity is considerable.

VII. CONCLUSION

In this paper, we considered the rank property of the optimal input covariance matrix under the average power constraint for a general MIMO Gaussian wiretap channel, where each node has an arbitrary number of antennas. We obtained necessary and sufficient constraints on the MIMO wiretap channel parameters
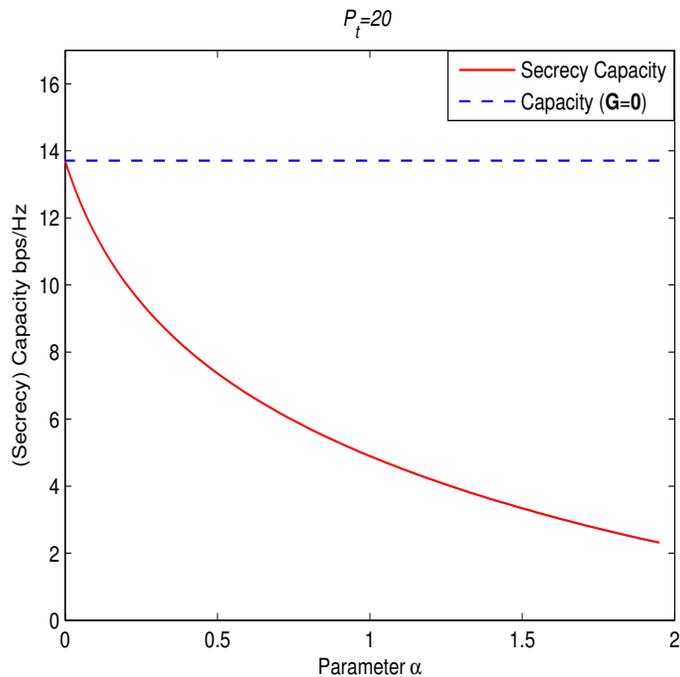
Fig. 4: Secrecy Capacity versus α for $P_t = 20$. Solid curve represents secrecy capacity and dashed curve indicates the point to point capacity.

such that the optimal input covariance matrix is full-rank, and we presented a method for characterizing the resulting covariance matrix as well.

APPENDIX A

PROOF OF LEMMA 4

Define $\bar{\mathbf{S}} = (\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G})^{-1}$ and apply the generalized eigenvalue decomposition on the pencil $(\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \ , \ \bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I})$ to obtain the invertible generalized eigenvector matrix $\bar{\mathbf{C}}$ and the diagonal generalized eigenvalue matrix $\mathbf{\Lambda}_{\bar{\mathbf{s}}}$ as

$$\bar{\mathbf{C}}^H \left[\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \right] \bar{\mathbf{C}} = \mathbf{\Lambda}_{\bar{\mathbf{s}}} \tag{66}$$

$$\bar{\mathbf{C}}^H \left[\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \right] \bar{\mathbf{C}} = \mathbf{I} . \tag{67}$$

By subtracting (67) from (66), we have

$$\bar{\mathbf{C}}^H \bar{\mathbf{C}} = \mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I} . \tag{68}$$

Note that from Lemma 3, we have $\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I} \succ \underline{0}$. Thus, $\bar{\mathbf{C}}$ must be of the form [14]

$$\bar{\mathbf{C}} = \mathbf{\Phi}_{\bar{\mathbf{s}}} \left(\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I} \right)^{\frac{1}{2}}, \tag{69}$$

where $\mathbf{\Phi}_{\bar{\mathbf{s}}}$ is an unknown unitary matrix. In the following, as we continue the proof, $\mathbf{\Phi}_{\bar{\mathbf{s}}}$ is also characterized.

By replacing (69) in (66) and (67), it is revealed that the unitary matrix $\mathbf{\Phi}_{\bar{\mathbf{s}}}$ represents the common set of eigenvectors for the matrices $\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I}$ and $\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I}$, and thus both matrices commute. In particular,

$$\mathbf{\Phi}_{\bar{\mathbf{s}}}^H \left[\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \right] \mathbf{\Phi}_{\bar{\mathbf{s}}} = \mathbf{\Lambda}_{\bar{\mathbf{s}}} (\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I})^{-1} = \mathbf{I} + (\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I})^{-1} \tag{70}$$

$$\mathbf{\Phi}_{\bar{\mathbf{s}}}^H \left[\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \right] \mathbf{\Phi}_{\bar{\mathbf{s}}} = (\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I})^{-1}. \tag{71}$$

Defining $\mathbf{D} = (\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I})^{-1}$, from (70)-(71) and via straightforward computation, we have

$$\mathbf{H}^H \mathbf{H} = \bar{\mathbf{S}}^{-\frac{1}{2}} \mathbf{\Phi}_{\bar{\mathbf{s}}} \mathbf{D} \mathbf{\Phi}_{\bar{\mathbf{s}}}^H \bar{\mathbf{S}}^{-\frac{1}{2}} \tag{72}$$

$$\mathbf{G}^H \mathbf{G} = \bar{\mathbf{S}}^{-\frac{1}{2}} \mathbf{\Phi}_{\bar{\mathbf{s}}} (\mathbf{D} - \mathbf{I}) \mathbf{\Phi}_{\bar{\mathbf{s}}}^H \bar{\mathbf{S}}^{-\frac{1}{2}}, \tag{73}$$

which proves (36) and (37). Substituting (68) in (67), we also have

$$\mathbf{I} = \bar{\mathbf{C}}^H \left[\bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} + \mathbf{I} \right] \bar{\mathbf{C}}$$

$$= \bar{\mathbf{C}}^H \bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} \bar{\mathbf{C}} + \mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I},$$

or equivalently

$$2\mathbf{I} - \mathbf{\Lambda}_{\bar{\mathbf{s}}} = \bar{\mathbf{C}}^H \bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} \bar{\mathbf{C}}.$$

Remark 5. Since $\bar{\mathbf{C}}^H \bar{\mathbf{S}}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \bar{\mathbf{S}}^{\frac{1}{2}} \bar{\mathbf{C}} \succeq \underline{0}$, it results that $2\mathbf{I} - \mathbf{\Lambda}_{\bar{\mathbf{s}}} \succeq \underline{0}$. Equivalently, by defining $\mathbf{D} = (\mathbf{\Lambda}_{\bar{\mathbf{s}}} - \mathbf{I})^{-1} \succ \underline{0}$, we have

$$\mathbf{I} - \mathbf{D}^{-1} \succeq \underline{0}$$
$$\mathbf{D} - \mathbf{I} \succeq \underline{0}. \tag{74}$$

Note that from (73), if $\mathbf{G}^H \mathbf{G} \succ \underline{0}$, then $\mathbf{D} - \mathbf{I} \succ \underline{0}$ and vice versa. As we will observe in Theorem 2, to have a full-rank optimal input covariance matrix $\mathbf{Q}^*$, having a full-rank $\mathbf{G}^H \mathbf{G}$ is not required. While we assume throughout the paper and without loss of generality that the diagonal matrix $\mathbf{D} - \mathbf{I}$ is invertible, for the case of rank deficient $\mathbf{G}^H \mathbf{G}$ one can follow the calculations in this paper assuming $\epsilon > 0$ for zero-diagonal elements of $\mathbf{D} - \mathbf{I}$ and letting $\epsilon \downarrow 0$ at the end (see Lemma 7).

APPENDIX B

PROOF OF THEOREM 3

We want to obtain $\mathbf{Q}_d^*$, the optimal input covariance matrix that attains the secrecy capacity for the case of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} \succeq \underline{0}$. We note that $\text{rank}(\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}) = m < n_t$. Hence, from Theorems 1 and 2, $\mathbf{Q}_d^*$ is rank-*deficient*.

The right hand side of (25) can be rewritten as

$$
\begin{aligned}
R(\mathbf{Q}) &= \log|\mathbf{I} + \mathbf{H}^H\mathbf{H}\mathbf{Q}| - \log|\mathbf{I} + \mathbf{G}^H\mathbf{G}\mathbf{Q}| \\
&= \log\left|\left(\mathbf{I} + \mathbf{H}^H\mathbf{H}\mathbf{Q}\right)\left(\mathbf{I} + \mathbf{G}^H\mathbf{G}\mathbf{Q}\right)^{-1}\right| \\
&= \log\left|\mathbf{I} + \left(\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}\right)\mathbf{Q}\left(\mathbf{I} + \mathbf{G}^H\mathbf{G}\mathbf{Q}\right)^{-1}\right|,
\end{aligned}
\tag{75}
$$

where Eq. (75) is obtained from the matrix inversion lemma [14] $(\mathbf{I} + \mathbf{A})^{-1} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{A})^{-1}$. Let the eigenvalue decomposition of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$ to be denoted as

$$
\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G} = \boldsymbol{\Psi} \begin{bmatrix} \boldsymbol{\Lambda}_m & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \boldsymbol{\Psi}^H,
\tag{76}
$$

where $\boldsymbol{\Lambda}_m \succeq \underline{0}$ is a diagonal matrix of size $m \times m$. Using (76) in (75), we have

$$
\begin{aligned}
R(\mathbf{Q}) &= \log\left|\mathbf{I} + \boldsymbol{\Psi}\begin{bmatrix} \boldsymbol{\Lambda}_m & 0 \\ \underline{0} & \underline{0} \end{bmatrix}\boldsymbol{\Psi}^H\mathbf{Q}\left(\mathbf{I} + \mathbf{G}^H\mathbf{G}\mathbf{Q}\right)^{-1}\right| \\
&= \log\left|\mathbf{I} + \begin{bmatrix} \boldsymbol{\Lambda}_m & 0 \\ \underline{0} & \underline{0} \end{bmatrix}\boldsymbol{\Psi}^H\mathbf{Q}\boldsymbol{\Psi}(\mathbf{I} + \boldsymbol{\Psi}^H\mathbf{G}^H\mathbf{G}\boldsymbol{\Psi}\,\boldsymbol{\Psi}^H\mathbf{Q}\boldsymbol{\Psi})^{-1}\right|,
\end{aligned}
\tag{77}
$$

where in obtaining (77) we have used the facts that $\boldsymbol{\Psi}^H\boldsymbol{\Psi} = \boldsymbol{\Psi}\boldsymbol{\Psi}^H = \mathbf{I}$ and $|\mathbf{I} + \mathbf{AB}| = |\mathbf{I} + \mathbf{BA}|$.

Define $\bar{\mathbf{Q}} = \boldsymbol{\Psi}^H\mathbf{Q}\boldsymbol{\Psi}$ and $\bar{\mathbf{G}} = \mathbf{G}\boldsymbol{\Psi}$, so that the optimization problem in (25) can be rewritten as

$$
\mathcal{C}_{sec}(P_t) = \max_{\bar{\mathbf{Q}}\succeq\underline{0},\,\text{Tr}(\bar{\mathbf{Q}})=P_t} R(\bar{\mathbf{Q}}),
$$

where

$$
R(\bar{\mathbf{Q}}) = \log\left|\mathbf{I} + \begin{bmatrix} \boldsymbol{\Lambda}_m & 0 \\ \underline{0} & \underline{0} \end{bmatrix}\bar{\mathbf{Q}}\left(\mathbf{I} + \bar{\mathbf{G}}^H\bar{\mathbf{G}}\,\bar{\mathbf{Q}}\right)^{-1}\right|.
\tag{78}
$$

From right-hand side of (78), we see that the optimal $\bar{\mathbf{Q}}$ is of the form

$$
\bar{\mathbf{Q}} = \begin{bmatrix} \bar{\mathbf{Q}}_m & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix},
\tag{79}
$$

where $\bar{\mathbf{Q}}_m \succeq \underline{0}$ is of size $m \times m$. Write $\bar{\mathbf{G}}^H \bar{\mathbf{G}}$ as

$$\bar{\mathbf{G}}^H \bar{\mathbf{G}} = \begin{bmatrix} \mathbf{J}_1 & \mathbf{J}_2 \\ \mathbf{J}_2^H & \mathbf{J}_3 \end{bmatrix} \tag{80}$$

where $\mathbf{J}_1$, $\mathbf{J}_2$ and $\mathbf{J}_3$ are of dimensions $m \times m$, $m \times (n_t - m)$ and $(n_t - m) \times (n_t - m)$, respectively. By substituting (79) and (80) into (78), we obtain

$$\begin{aligned} R(\bar{\mathbf{Q}}) &= \log \left| \mathbf{I} + \begin{bmatrix} \mathbf{\Lambda}_m \bar{\mathbf{Q}}_m & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} \mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m & \underline{0} \\ \mathbf{J}_2^H \bar{\mathbf{Q}}_m & \mathbf{I} \end{bmatrix}^{-1} \right| \\ &= \log \left| \mathbf{I} + \begin{bmatrix} \mathbf{\Lambda}_m \bar{\mathbf{Q}}_m & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} (\mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m)^{-1} & \underline{0} \\ -\mathbf{J}_2^H \bar{\mathbf{Q}}_m (\mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m)^{-1} & \mathbf{I} \end{bmatrix} \right| \\ &= \log \left| \mathbf{I} + \begin{bmatrix} \mathbf{\Lambda}_m \bar{\mathbf{Q}}_m (\mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \right| \\ &= \log \left| \begin{bmatrix} \mathbf{I} + \mathbf{\Lambda}_m \bar{\mathbf{Q}}_m (\mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m)^{-1} & \underline{0} \\ \underline{0} & \mathbf{I} \end{bmatrix} \right| \\ &= \log \left| \mathbf{I} + \mathbf{\Lambda}_m \bar{\mathbf{Q}}_m (\mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m)^{-1} \right| \\ &= \log \left| \mathbf{I} + (\mathbf{\Lambda}_m + \mathbf{J}_1) \bar{\mathbf{Q}}_m \right| - \log \left| \mathbf{I} + \mathbf{J}_1 \bar{\mathbf{Q}}_m \right| = R(\bar{\mathbf{Q}}_m). \end{aligned} \tag{81}$$

Using (81), the secrecy capacity is given by

$$\mathcal{C}_{sec}(P_t) = \max_{\bar{\mathbf{Q}}_m \succeq \underline{0},\ \mathrm{Tr}(\bar{\mathbf{Q}}_m) = P_t} R(\bar{\mathbf{Q}}_m) . \tag{82}$$

Problem (82) shows that the secrecy capacity of a wiretap channel with $\mathbf{H}^H \mathbf{H} - \mathbf{G}^H \mathbf{G} \succeq \underline{0}$ is equal to the secrecy capacity of an equivalent wiretap channel with

$$\mathbf{H}_{eq}^H \mathbf{H}_{eq} = \mathbf{\Lambda}_m + \mathbf{J}_1 \tag{83}$$

$$\mathbf{G}_{eq}^H \mathbf{G}_{eq} = \mathbf{J}_1 , \tag{84}$$

where $\mathbf{\Lambda}_m$ and $\mathbf{J}_1$ are respectively given by (76) and (80). It should also be noted that for the equivalent channel, $\mathbf{H}_{eq}^H \mathbf{H}_{eq} - \mathbf{G}_{eq}^H \mathbf{G}_{eq} = \mathbf{\Lambda}_m \succ \underline{0}$. Thus, the optimal $\bar{\mathbf{Q}}_m^*$ can be computed using Theorem 2, as long as the equivalent channel satisfies the second condition in Theorem 2. Finally, by substituting $\bar{\mathbf{Q}}_m^*$ back into (79) we obtain

$$\mathbf{Q}_d^* = \mathbf{\Psi} \begin{bmatrix} \bar{\mathbf{Q}}_m^* & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{\Psi}^H , \tag{85}$$

which completes the proof.

REFERENCES

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

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. 24, pp. 451-456, Jul. 1978.

[3] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. IEEE Int. Symp. Information Theory* Toronto, ON, Canada, Jul. 2008, pp. 524-528.

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

[5] J. Li and A. P. Petropulu, "Transmitter optimization for achieving secrecy capacity in Gaussian MIMO wiretap channels," submitted to *IEEE Trans. Info. Theory*, Available [online]: http://arxiv.org/PS cache/arxiv/pdf/0909/0909.2622v1.pdf.

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

[7] S. Shafiee and S. Ulukus, "Towards the Secrecy Capacity of the Gaussian MIMO Wire-Tap Channel: The 2-2-1 Channel," *IEEE Trans. on Inf. Theory*, vol. 55, no. 9, Sep. 2009.

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

[9] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "A MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, Article ID 370970, 8 pages, 2009.

[10] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multipleoutput broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936-3964, 2006.

[11] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493-2512, June 2008.

[12] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235-1249, Mar. 2009.

[13] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215-4227, 2010.

[14] R. A. Horn and C. R. Johnson, *Matrix Analysis*, University Press, Cambridge, UK, 1999.

[15] C. F. V. Loan, "Generalizing the Singular Value Decomposition," *SIAM Journal on Num. Analysis*, vol. 13, no. 1, pp. 76-83, Mar. 1973.