Towards the Secrecy Capacity of the Gaussian MIMO Wire-Tap Channel: The 2-2-1 Channel

Shabnam Shafiee, Nan Liu, *Member, IEEE*, and Sennur Ulukus, *Member, IEEE*

***Abstract*—We find the secrecy capacity of the 2-2-1 Gaussian MIMO wiretap channel, which consists of a transmitter and a receiver with two antennas each, and an eavesdropper with a single antenna. We determine the secrecy capacity of this channel by proposing an achievable scheme and then developing a tight upper bound that meets the proposed achievable secrecy rate. We show that, for this channel, Gaussian signalling in the form of beam-forming is optimal, and no pre-processing of information is necessary.**

***Index Terms*—Information-theoretic security, multiple-input multiple-output (MIMO), multiple antennas, secrecy capacity, wiretap channel.**

I. Introduction

THE inherent openness of wireless communications makes it vulnerable to eavesdropping and jamming attacks. This vulnerability has to be addressed through secure communications. The eavesdropping attack was first studied by Wyner in [1], where he considers a single-user wiretap channel. The measure of secrecy is the message equivocation rate at the wire-tapper, which is defined as the entropy rate of the message at the wire-tapper, given the wire-tapper's observation. Wyner models the wire-tapper's channel as a degraded version of the channel from the transmitter to the legitimate receiver, which is a reasonable assumption in a wired channel. For this channel, Wyner identifies the rate-equivocation region and therefore, the secrecy capacity. Wyner's result was extended to the Gaussian wiretap channel in [2], and it was shown that Gaussian signalling is optimal. The secrecy capacity was found to be the difference between the capacities of the main and the eavesdropping channels.

Csiszár and Körner [3] studied the general, i.e., not necessarily degraded, single-transmitter, single-receiver, single-eavesdropper, discrete memoryless channel with secrecy constraints, and found an expression for the secrecy capacity, in the form of the maximization of the difference between two mutual informations involving an auxiliary random variable. The auxiliary random variable is interpreted as performing pre-processing on the information. The explicit calculation of the secrecy capacity for a given channel requires the solution of this maximization problem in terms of the joint distribution of the auxiliary random variable and the channel input.

The use of multiple transmit and receive antennas has been shown to increase the achievable rates when there are no secrecy constraints [4]. The Gaussian multiple-input multiple-output (MIMO) wiretap channel is a special case of the single-transmitter, single-receiver, single-eavesdropper wiretap channel. Since the Gaussian MIMO channel is not degraded in general, finding its secrecy capacity involves identifying the optimum joint distribution of the auxiliary random variable representing pre-processing and the channel input in the Csiszár-Körner formula. However, solving this optimization problem directly for non-degraded channels is difficult, forcing researchers typically to follow a two-step solution, where in the first step a feasible solution is identified (an achievable scheme), and in the second step a tight upper bound that meets this feasible solution is developed (tight converse).

The first paper studying secrecy in MIMO communications is [5], which proposes an achievable scheme, where the transmitter uses its multiple transmit antennas to transmit only in the null space of the eavesdropper's channel, thereby preventing any eavesdropping. Reference [6] studies the Gaussian single-input multiple-output (SIMO) wiretap channel, and shows that it is equivalent to a scalar Gaussian channel, and gives the secrecy capacity using the results of [2]. An achievable scheme has been proposed for the Gaussian multiple-input single-output (MISO) wiretap channel in [7], and independently and concurrently in [8]. In both of these papers, the achievable secrecy rate is obtained by restricting the channel input to be Gaussian, with no pre-processing of information. The secrecy rate found in [7], [8] is shown to be the secrecy capacity of the Gaussian MISO wiretap channel in [9], [10]. Further, [9], [10] allow the eavesdropper to have multiple antennas (MISOME).

In all of the above papers, the secrecy capacity of MIMO communications is specified only in the cases where the receiver has a single antenna. The next step towards finding the secrecy capacity of the general Gaussian MIMO channel is to consider multiple antennas at the receiver. In this paper, we consider a MIMO channel where both the transmitter and the receiver have multiple antennas. More specifically, we focus on a simple special case where both the transmitter and the receiver have two antennas each, and the eavesdropper has a single antenna, hence we call this channel the 2-2-1 MIMO wiretap channel. We find the secrecy capacity in two steps: we first propose an

Manuscript received September 21, 2007; revised December 23, 2008. Current version published August 19, 2009. This work was supported by the NSF under Grants CCR 03-11311, CCF 04-47613, and CCF 05-14846. The material of this paper was presented in part at the 3rd International Symposium on Communications, Control and Signal Processing, St. Julians, Malta, March 2008.

S. Shafiee was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA. She is now with Perinatronics Medical Systems, Millersville, MD 21108-2519 USA (e-mail: sshafiee@perinatronics.com).

N. Liu was with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA. She is now with the School of Information Science and Engineering, Southeast University, Nanjing, China (e-mail: nkancy@umd.edu).

S. Ulukus is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742 USA (e-mail: ulukus@umd.edu).

Communicated by A. J. Grant, Associate Editor for Communications.

Digital Object Identifier 10.1109/TIT.2009.2025549

achievable scheme, which is a Gaussian signalling scheme with no pre-processing of information, and then, we develop a tight upper bound that meets the rate achieved with our proposed signalling scheme.

We first show that the optimal Gaussian signalling scheme has a unit-rank transmit covariance matrix, hence with Gaussian signalling, beam-forming is optimal. The transmitter beam-forms in a direction that is as orthogonal to the direction of the eavesdropper, and as close to the two directions of the receiver as possible. Then, we develop an upper bound by considering a channel where the eavesdropper's signal is given to the receiver. The secrecy capacity of this channel is an upper bound to the secrecy capacity of the original channel. In addition, this channel is degraded, and no pre-processing of information is needed. Furthermore, Gaussian signalling is optimal for this channel. We further tighten this bound by allowing correlation between the additive noises of the receiver and the eavesdropper. For a certain such correlation, we prove that the optimal Gaussian signalling is unit-rank in this upper bound also. We then evaluate our upper bound and show that it meets the rate achievable with our proposed signalling scheme. In this 2-2-1 system, the fact that both in our achievable scheme and in our upper bound, the optimal transmit covariance matrices turn out to be unit-rank, proves to be crucial in enabling us to characterize the lower and upper bounds explicitly and showing that they are equal.

Secure communications in multi-user networks, e.g., multiple access channel [11]–[15], broadcast channel [16], relay channel [17], [18], interference channel [19], and two-way channel [20], and in fading channels [7], [21]–[25] have been considered recently.

We use the following notations throughout this paper: Bold face lower and upper case letters are used to represent vectors and matrices, respectively. $\boldsymbol{x}^T$ and $||\boldsymbol{x}||$ denote the transpose and the Euclidean norm of the vector $\boldsymbol{x}$, respectively. $\mathrm{tr}(\boldsymbol{X})$ and $|\boldsymbol{X}|$ denote the trace and the determinant of the square matrix $\boldsymbol{X}$, respectively. The notation $[x]^+$ means $\max(x, 0)$. Whether a variable is deterministic or random will be clear from the context.

II. System Model

The 2-2-1 Gaussian MIMO wiretap channel is characterized by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}_y \tag{1}$$

$$z = \boldsymbol{g}^T\boldsymbol{x} + n_z \tag{2}$$

where $\boldsymbol{x}$ is the transmitted signal, and $\boldsymbol{y}$, z are the received signals at the legitimate user and the eavesdropper, respectively. $\boldsymbol{n}_y$ is a Gaussian random vector with zero-mean and identity covariance matrix, while n_z is a Gaussian random variable with zero-mean and unit-variance. $\boldsymbol{n}_y$, n_z are assumed to be independent. The transmitted signal satisfies an average power constraint

$$\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{x}_i^T\boldsymbol{x}_i \le P. \tag{3}$$

The secrecy capacity $C(P)$ is defined as the maximum number of bits that can be correctly transmitted to the intended receiver while the eavesdropper is essentially no better informed about the transmitted information after observing the received signal than it was before [2]. The secrecy capacity of the 2-2-1 Gaussian MIMO wiretap channel as defined in (1) and (2) is given by the following general expression [3]:

$$\max_{p(\boldsymbol{u},\boldsymbol{x})}[I(\boldsymbol{u};\boldsymbol{y}) - I(\boldsymbol{u};z)]^+ \tag{4}$$

where the maximum is over all distributions that satisfy $\boldsymbol{u} \to \boldsymbol{x} \to \boldsymbol{y}z$. The problem studied in this paper is to find the maximizing $p(\boldsymbol{u},\boldsymbol{x})$, and provide an explicit formula for the secrecy capacity. The expression in (4) shows that the secrecy capacity only depends on the marginals of the two channels $p(\boldsymbol{y}|\boldsymbol{x})$ and $p(z\,|\,\boldsymbol{x})$, and is independent of the correlation between the two channels, i.e., the correlation between $\boldsymbol{n}_y$ and n_z.

When $\boldsymbol{H}$ is not full-rank, by performing singular value decomposition (SVD) on $\boldsymbol{H}$ and obtaining an equivalent channel by rotation, it can be shown that the system is equivalent to a 2-1-1 system, whose secrecy capacity has been found in [9], [10]. Therefore, without loss of generality, for the rest of the paper, we assume that $\boldsymbol{H}$ is full-rank, and hence is invertible. When

$$||\boldsymbol{H}^{-T}\boldsymbol{g}|| \le 1 \tag{5}$$

z can be written as a noisy version of $\boldsymbol{y}$, i.e., $\boldsymbol{r}^T\boldsymbol{y} + n$, which means that the channel is degraded. In this case, no pre-processing of information is necessary [3], and also it can be shown that Gaussian signalling is optimal. Thus, in this paper, we concentrate on the more interesting and difficult case where $\boldsymbol{H}$ is full-rank and satisfies

$$||\boldsymbol{H}^{-T}\boldsymbol{g}|| > 1. \tag{6}$$

III. An Achievable Scheme

By [3], the following secrecy rate is achievable

$$[I(\boldsymbol{u};\boldsymbol{y}) - I(\boldsymbol{u};z)]^+ \tag{7}$$

where $\boldsymbol{u} \to \boldsymbol{x} \to \boldsymbol{y}z$. By taking $\boldsymbol{u} = \boldsymbol{x}$ and constraining the input signal $\boldsymbol{x}$ to be Gaussian with covariance matrix $\boldsymbol{S}$ such that $\mathrm{tr}(\boldsymbol{S}) \le P$, the following secrecy rate is achievable

$$\left[\frac{1}{2}\log|\boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T| - \frac{1}{2}\log(1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})\right]^+. \tag{8}$$

Thus, the following secrecy rate is achievable

$$\max_{\boldsymbol{S}\succeq\boldsymbol{0}:\mathrm{tr}(\boldsymbol{S})\le P}\frac{1}{2}\log|\boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T| - \frac{1}{2}\log(1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}). \tag{9}$$

Here, we are able to remove the $[\cdot]^+$ sign because the maximum value in (9) is always strictly positive. The reason is as follows: By picking $\boldsymbol{S} = P\boldsymbol{g}^\perp(\boldsymbol{g}^\perp)^T$, where $\boldsymbol{g}^\perp$ is the unit-norm vector that is orthogonal to $\boldsymbol{g}$, an achievable secrecy rate is

$$\frac{1}{2}\log(1 + P||\boldsymbol{H}\boldsymbol{g}^\perp||^2). \tag{10}$$

Based on the fact that $\boldsymbol{H}$ is full rank, i.e., $\boldsymbol{H}\boldsymbol{g}^{\perp} \neq \boldsymbol{0}$, the secrecy rate in (10) is strictly positive. Since the secrecy rate in (9) is the maximum over all $\boldsymbol{S}$ satisfying $\mathrm{tr}(\boldsymbol{S}) \leq P$, we conclude that it must be strictly positive as well.

Ignoring the $1/2$, we may rewrite the cost function in (9) as

$$\log|\boldsymbol{I}+\boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T| - \log(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}) = \log|\boldsymbol{I}+\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{S}| - \log(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}). \quad (11)$$

We first use the following lemma to show that the $\boldsymbol{S}$ that maximizes (9) is unit-rank.

Lemma 1: If $\boldsymbol{D}$ is a 2×2 invertible matrix that satisfies

$$\boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g} \geq 1 \quad (12)$$

then the optimal $\boldsymbol{S}$ that solves the following optimization problem

$$\max_{\boldsymbol{S}\succeq\boldsymbol{0},\mathrm{tr}(\boldsymbol{S})\leq P} \log|\boldsymbol{I}+\boldsymbol{D}\boldsymbol{S}| - \log(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}) \quad (13)$$

is unit-rank.

Proof: The KKT necessary conditions for the optimization problem in (13) are

$$\boldsymbol{S}^* \succeq \boldsymbol{0} \quad (14)$$
$$\mathrm{tr}(\boldsymbol{S}^*) \leq P \quad (15)$$
$$\boldsymbol{C} \succeq \boldsymbol{0} \quad (16)$$
$$\lambda \geq 0 \quad (17)$$
$$\lambda(\mathrm{tr}(\boldsymbol{S}^*) - P) = 0 \quad (18)$$
$$\boldsymbol{C}\boldsymbol{S}^* = \boldsymbol{0} \quad (19)$$
$$-(\boldsymbol{I}+\boldsymbol{D}\boldsymbol{S}^*)^{-1}\boldsymbol{D} + \frac{1}{1+\boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}}\boldsymbol{g}\boldsymbol{g}^T - \boldsymbol{C} + \lambda\boldsymbol{I} = \boldsymbol{0} \quad (20)$$

We will prove the claim by contradiction. Assume that the optimal $\boldsymbol{S}$ is full-rank. Then, from (19), it follows that $\boldsymbol{C} = \boldsymbol{0}$, i.e., (20) becomes

$$(\boldsymbol{I}+\boldsymbol{D}\boldsymbol{S}^*)^{-1}\boldsymbol{D} = \frac{1}{1+\boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}}\boldsymbol{g}\boldsymbol{g}^T + \lambda\boldsymbol{I}. \quad (21)$$

Since $\boldsymbol{D}$ is invertible

$$(\boldsymbol{I}+\boldsymbol{D}\boldsymbol{S}^*)^{-1} = \frac{1}{1+\boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}}\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{D}^{-1} + \lambda\boldsymbol{D}^{-1}. \quad (22)$$

Using the matrix inversion lemma [26, p. 19], we have

$$\boldsymbol{I}+\boldsymbol{D}\boldsymbol{S}^* = \frac{1}{\lambda}\boldsymbol{D} - \frac{1}{\lambda^2+\lambda^2\boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}+\lambda||\boldsymbol{g}||^2}\boldsymbol{D}\boldsymbol{g}\boldsymbol{g}^T \quad (23)$$

i.e.,

$$\boldsymbol{S}^* = \frac{1}{\lambda}\boldsymbol{I} - \frac{1}{\lambda^2+\lambda^2\boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}+\lambda||\boldsymbol{g}||^2}\boldsymbol{g}\boldsymbol{g}^T - \boldsymbol{D}^{-1}. \quad (24)$$

We multiply both sides of (24) with $\boldsymbol{g}^T$ on the left and $\boldsymbol{g}$ on the right. Let us define $\gamma = \boldsymbol{g}^T\boldsymbol{S}^*\boldsymbol{g}$, which is a nonnegative real number. Then, we have

$$\gamma = \frac{||\boldsymbol{g}||^2}{\lambda} - \frac{||\boldsymbol{g}||^4}{\lambda^2+\lambda^2\gamma+\lambda||\boldsymbol{g}||^2} - \boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g} \quad (25)$$

i.e., we have

$$\gamma^2 + (1+\boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g})\gamma + \boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g} + \frac{||\boldsymbol{g}||^2}{\lambda}\left(\boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g} - 1\right) = 0. \quad (26)$$

Because $\boldsymbol{g}^T\boldsymbol{D}^{-1}\boldsymbol{g} - 1 \geq 0$, all coefficients in the second order equation in (26) are positive. Hence, the second-order equation has no non-negative roots, i.e., it either has no real roots, or it has two negative roots. Thus, we arrive at a contradiction. Therefore, $\boldsymbol{C}$ cannot be equal to $\boldsymbol{0}$, and consequently, $\boldsymbol{S}$ cannot be full-rank, and it has to be unit-rank. □

Since $\boldsymbol{H}^T\boldsymbol{H}$ is invertible and satisfies (6), $\boldsymbol{D} = \boldsymbol{H}^T\boldsymbol{H}$ satisfies the condition of Lemma 1. Hence, the optimal $\boldsymbol{S}$ for the optimization problem in (9) is unit-rank.

Given that the optimal $\boldsymbol{S}$ is unit-rank, it can be written as

$$\boldsymbol{S} = P\boldsymbol{q}\boldsymbol{q}^T. \quad (27)$$

The corresponding achievable secrecy rate is

$$R = \frac{1}{2}\log|\boldsymbol{I}+P\boldsymbol{H}\boldsymbol{q}\boldsymbol{q}^T\boldsymbol{H}^T| - \frac{1}{2}\log(1+P\boldsymbol{g}^T\boldsymbol{q}\boldsymbol{q}^T\boldsymbol{g}) \quad (28)$$
$$= \frac{1}{2}\log\frac{\boldsymbol{q}^T(\boldsymbol{I}+P\boldsymbol{H}^T\boldsymbol{H})\boldsymbol{q}}{\boldsymbol{q}^T(\boldsymbol{I}+P\boldsymbol{g}\boldsymbol{g}^T)\boldsymbol{q}} \quad (29)$$

where (29) is now in the Rayleigh quotient [26, p. 176] form and the optimal achievable $\boldsymbol{q}$, which we will call $\boldsymbol{q}_a$, is

$$\boldsymbol{q}_a = \frac{\boldsymbol{B}^{-1/2}\boldsymbol{w}_a}{||\boldsymbol{B}^{-1/2}\boldsymbol{w}_a||} \quad (30)$$

where $\boldsymbol{w}_a$ is the eigenvector that corresponds to the largest eigenvalue of $\boldsymbol{B}^{-1/2}\boldsymbol{A}\boldsymbol{B}^{-1/2}$ with

$$\boldsymbol{A} = \boldsymbol{I}+P\boldsymbol{H}^T\boldsymbol{H} \quad (31)$$
$$\boldsymbol{B} = \boldsymbol{I}+P\boldsymbol{g}\boldsymbol{g}^T. \quad (32)$$

In other words, $\boldsymbol{q}_a$ is the unit-norm eigenvector that satisfies

$$(\boldsymbol{I}+P\boldsymbol{g}\boldsymbol{g}^T)^{-1}(\boldsymbol{I}+P\boldsymbol{H}^T\boldsymbol{H})\boldsymbol{q}_a = \lambda_1\boldsymbol{q}_a \quad (33)$$

where λ_1 is the largest eigenvalue of the matrix

$$(\boldsymbol{I}+P\boldsymbol{g}\boldsymbol{g}^T)^{-1/2}(\boldsymbol{I}+P\boldsymbol{H}^T\boldsymbol{H})(\boldsymbol{I}+P\boldsymbol{g}\boldsymbol{g}^T)^{-1/2}. \quad (34)$$

Written explicitly, the achievable secrecy rate is

$$\frac{1}{2}\log\left(\frac{1+P\boldsymbol{q}_a^T\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{q}_a}{1+P\boldsymbol{q}_a^T\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{q}_a}\right) = \frac{1}{2}\log\lambda_1. \quad (35)$$

Based on the argument made after (10), we have

$$\frac{1}{2}\log\lambda_1 \geq \frac{1}{2}\log(1+P||\boldsymbol{H}\boldsymbol{g}^{\perp}||^2) > 0 \quad (36)$$

which also means that

$$\lambda_1 > 1. \quad (37)$$

IV. A Tight Upper Bound

The following theorem provides an upper bound on the secrecy capacity of the wiretap channel described in (1) and (2).

Theorem 1: An upper bound on the secrecy capacity of the wiretap channel described in (1) and (2) is

$$\max_{\boldsymbol{S}\succeq\boldsymbol{0},\mathrm{tr}(\boldsymbol{S})\leq P} U(\boldsymbol{S},\boldsymbol{a}) \tag{38}$$

for any $\boldsymbol{a}$ with $||\boldsymbol{a}|| < 1$, where $U(\boldsymbol{S},\boldsymbol{a})$ is defined as

$$U(\boldsymbol{S},\boldsymbol{a}) = \frac{1}{2}\log\frac{|\boldsymbol{I}+\boldsymbol{N}^{-1}\bar{\boldsymbol{H}}\boldsymbol{S}\bar{\boldsymbol{H}}^T|}{(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})} \tag{39}$$

with $\boldsymbol{N}$ defined as

$$\boldsymbol{N} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{a} \\ \boldsymbol{a}^T & 1 \end{bmatrix} \tag{40}$$

and $\bar{\boldsymbol{H}}$ defined as

$$\bar{\boldsymbol{H}} = \begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{g}^T \end{bmatrix}. \tag{41}$$

The Proof of Theorem 1 is provided in the Appendix. Intuitively, this upper bound is obtained by considering the secrecy capacity of a new channel where the legitimate receiver also has access to the eavesdropper's signal. Since the legitimate user is more capable in the new channel, the secrecy capacity of the new channel will serve as an upper bound on the secrecy capacity of the original channel. The new channel is degraded, and therefore the secrecy capacity is easier to obtain.

The vector $\boldsymbol{a}$ introduced in Theorem 1 is the correlation between the Gaussian noises at the legitimate user and the eavesdropper, i.e.,

$$\boldsymbol{a} = E[\boldsymbol{n}_y n_z]. \tag{42}$$

We note that $\boldsymbol{a}$ thus defined has to satisfy $||\boldsymbol{a}|| \leq 1$ for $\boldsymbol{N}$ in (40) to be positive semi-definite. Introducing correlation between $\boldsymbol{n}_y$ and n_z does not change the secrecy capacity of the channel as can be seen from (4), but changes the upper bound in (38). In fact, (38) remains a valid upper bound for any $\boldsymbol{a}$, with $||\boldsymbol{a}|| < 1$. Thus, we will smartly pick an $\boldsymbol{a}$ vector, and show that the upper bound with this $\boldsymbol{a}$ vector is in fact tight, to establish the secrecy capacity.

We rewrite $U(\boldsymbol{S},\boldsymbol{a})$ as

$$U(\boldsymbol{S},\boldsymbol{a}) = \frac{1}{2}\log\frac{|\boldsymbol{I}+\bar{\boldsymbol{H}}^T\boldsymbol{N}^{-1}\bar{\boldsymbol{H}}\boldsymbol{S}|}{(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})}. \tag{43}$$

By the definition of $\boldsymbol{N}$ in (40), we have

$$\boldsymbol{N}^{-1} = \begin{bmatrix} \boldsymbol{I}+\frac{1}{k}\boldsymbol{a}\boldsymbol{a}^T & -\frac{1}{k}\boldsymbol{a} \\ -\frac{1}{k}\boldsymbol{a}^T & \frac{1}{k} \end{bmatrix} \tag{44}$$

where $k = 1 - ||\boldsymbol{a}||^2$. Then

$$\begin{aligned}\bar{\boldsymbol{H}}^T\boldsymbol{N}^{-1}\bar{\boldsymbol{H}} &= \boldsymbol{H}^T\boldsymbol{H} + \frac{1}{k}\boldsymbol{H}^T\boldsymbol{a}\boldsymbol{a}^T\boldsymbol{H} - \frac{1}{k}\boldsymbol{g}\boldsymbol{a}^T\boldsymbol{H} - \frac{1}{k}\boldsymbol{H}^T\boldsymbol{a}\boldsymbol{g}^T + \frac{1}{k}\boldsymbol{g}\boldsymbol{g}^T \qquad (45)\\ &= \boldsymbol{H}^T\boldsymbol{H} + \frac{1}{k}(\boldsymbol{H}^T\boldsymbol{a}-\boldsymbol{g})(\boldsymbol{H}^T\boldsymbol{a}-\boldsymbol{g})^T \qquad (46)\end{aligned}$$

Let us define $\boldsymbol{A}(\boldsymbol{a})$ as

$$\boldsymbol{A}(\boldsymbol{a}) = \bar{\boldsymbol{H}}^T\boldsymbol{N}^{-1}\bar{\boldsymbol{H}} = \boldsymbol{H}^T\boldsymbol{H} + \frac{1}{k}(\boldsymbol{H}^T\boldsymbol{a}-\boldsymbol{g})(\boldsymbol{H}^T\boldsymbol{a}-\boldsymbol{g})^T. \tag{47}$$

Then, $U(\boldsymbol{S},\boldsymbol{a})$ in (43) is written as

$$U(\boldsymbol{S},\boldsymbol{a}) = \frac{1}{2}\log|\boldsymbol{I}+\boldsymbol{A}(\boldsymbol{a})\boldsymbol{S}| - \frac{1}{2}\log(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}). \tag{48}$$

Let us also define $\boldsymbol{q}_a^{\perp}$ to be the unit-norm vector that is orthogonal to $\boldsymbol{q}_a$, which is defined in (30).

We pick $\boldsymbol{a}$ to be of the form

$$\boldsymbol{a} = \boldsymbol{H}^{-T}(\alpha\boldsymbol{q}_a^{\perp} + \boldsymbol{g}) \tag{49}$$

for any real number α that makes $||\boldsymbol{a}|| < 1$. $\alpha = 0$ results in $\boldsymbol{a} = \boldsymbol{H}^{-T}\boldsymbol{g}$, which is a vector with norm greater than 1, and therefore, is not permissible.

Then, with this selection of $\boldsymbol{a}$, $\boldsymbol{A}(\boldsymbol{a})$ in (47) can be written as

$$\boldsymbol{A}(\boldsymbol{a}) = \bar{\boldsymbol{H}}^T\bar{\boldsymbol{H}} + \theta(\alpha)\boldsymbol{q}_a^{\perp}(\boldsymbol{q}_a^{\perp})^T \tag{50}$$

where $\theta(\alpha)$ is defined as

$$\begin{aligned}\theta(\alpha) &= \frac{\alpha^2}{1-\boldsymbol{a}^T\boldsymbol{a}} \qquad (51)\\ &= \frac{\alpha^2}{1-(\boldsymbol{H}^{-T}(\alpha\boldsymbol{q}_a^{\perp}+\boldsymbol{g}))^T(\boldsymbol{H}^{-T}(\alpha\boldsymbol{q}_a^{\perp}+\boldsymbol{g}))}. \qquad (52)\end{aligned}$$

Then, we have

$$\begin{aligned}\frac{1}{\theta(\alpha)} &= -(\boldsymbol{q}_a^{\perp})^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^{\perp} \\ &\quad - \frac{2\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^{\perp}}{\alpha} \\ &\quad - \frac{\boldsymbol{g}^T\left(\boldsymbol{H}^T\boldsymbol{H}\right)^{-1}\boldsymbol{g} - 1}{\alpha^2}\end{aligned} \tag{53}$$

This is a second-order polynomial in terms of $1/\alpha$, and it is easy to see that $1/\alpha^*$ maximizes $\theta(\alpha)$, with

$$\frac{1}{\alpha^*} = \frac{\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^{\perp}}{1-\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{g}}. \tag{54}$$

Finally, we call the $\boldsymbol{a}$ vector that we pick $\boldsymbol{a}^*$, which is given as

$$\boldsymbol{a}^* = \boldsymbol{H}^{-T}(\alpha^*\boldsymbol{q}_a^{\perp} + \boldsymbol{g}). \tag{55}$$

First, we will prove that $\boldsymbol{a}^*$ has norm no greater than 1. Let us define $\boldsymbol{a}_0$ to be

$$\boldsymbol{a}_0 = \frac{\boldsymbol{g}^T\boldsymbol{q}_a}{||\boldsymbol{H}\boldsymbol{q}_a||^2}\boldsymbol{H}\boldsymbol{q}_a \tag{56}$$

$\boldsymbol{a}_0$ satisfies the form of $\boldsymbol{a}$ in (49) because $\bar{\boldsymbol{H}}^T\boldsymbol{a}_0 - \boldsymbol{g}$ is orthogonal to $\boldsymbol{q}_a$, hence, it is along the direction of $\boldsymbol{q}_a^{\perp}$. Therefore, $\boldsymbol{a}_0$ must correspond to an α, which we call α_0. It can be seen that

$$||\boldsymbol{a}_0|| = \frac{|\boldsymbol{g}^T\boldsymbol{q}_a|}{||\boldsymbol{H}\boldsymbol{q}_a||} < 1 \tag{57}$$

because of (37) and the fact that $\boldsymbol{q}_a$ satisfies (35), i.e.,

$$1 < \lambda_1 = \frac{1 + P||\boldsymbol{H}\boldsymbol{q}_a||^2}{1 + P(\boldsymbol{g}^T\boldsymbol{q}_a)^2}. \tag{58}$$

Hence, (57) means that $\alpha_0 \neq 0$ and furthermore, we have

$$\theta(\alpha_0) = \frac{\alpha_0^2}{1 - \boldsymbol{a}_0^T\boldsymbol{a}_0} > 0. \tag{59}$$

Therefore, we have

$$\frac{1}{\theta(\alpha^*)} \overset{(a)}{\geq} \frac{1}{\theta(\alpha_0)} > 0 \tag{60}$$

where (a) follows because α^* maximizes $\frac{1}{\theta(\alpha^*)}$. Finally, (60) implies $||\boldsymbol{a}^*|| < 1$ because of (51).

Next, we will show that the optimal $\boldsymbol{S}$ for $\max U(\boldsymbol{S}, \boldsymbol{a}^*)$ in (38) is unit-rank. Since the upper bound and achievable scheme differ only in replacing $\boldsymbol{A}(\boldsymbol{a}^*)$ with $\boldsymbol{H}^T\boldsymbol{H}$, as shown in (9) and (48), we will use Lemma 1 again to show the optimality of unit-rank $\boldsymbol{S}$ in (38). Since $\boldsymbol{H}^T\boldsymbol{H}$ is invertible and $\theta(\alpha^*) > 0$, matrix $\boldsymbol{A}(\boldsymbol{a}^*)$, in the form of (50), is invertible. In addition, in order to use Lemma 1, we need $\boldsymbol{g}^T\boldsymbol{A}(\boldsymbol{a}^*)^{-1}\boldsymbol{g} \geq 1$. In the following, we will show that $\boldsymbol{g}^T\boldsymbol{A}(\boldsymbol{a}^*)^{-1}\boldsymbol{g} = 1$. Using the matrix inversion lemma [26, page 19] on (50), we have

$$\boldsymbol{A}(\boldsymbol{a}^*)^{-1} = (\boldsymbol{H}^T\boldsymbol{H})^{-1} - \frac{1}{\frac{1}{\theta(\alpha^*)} + (\boldsymbol{q}_a^\perp)^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^\perp} \times (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}. \tag{61}$$

Also, from (53) and (54), $1/\theta(\alpha^*)$ is equal to

$$\frac{1}{\theta(\alpha^*)} = -(\boldsymbol{q}_a^\perp)^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^\perp + \frac{(\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{q}_a^\perp)^2}{\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{g} - 1} \tag{62}$$

$$= -(\boldsymbol{q}_a^\perp)^T\left((\boldsymbol{H}^T\boldsymbol{H})^{-1} - \frac{(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{g}\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}}{\boldsymbol{g}^T(\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{g} - 1}\right)\boldsymbol{q}_a^\perp \tag{63}$$

$$= -(\boldsymbol{q}_a^\perp)^T(\boldsymbol{H}^T\boldsymbol{H} - \boldsymbol{g}\boldsymbol{g}^T)^{-1}\boldsymbol{q}_a^\perp. \tag{64}$$

Now, using straightforward algebra, starting from (61) and (62), it is easy to verify that

$$\boldsymbol{g}^T\boldsymbol{A}(\boldsymbol{a}^*)^{-1}\boldsymbol{g} = 1. \tag{65}$$

Thus, $\boldsymbol{D} = \boldsymbol{A}(\boldsymbol{a}^*)$ satisfies the conditions of Lemma 1, and therefore, $\arg\max U(\boldsymbol{S}, \boldsymbol{a}^*)$ is unit-rank.

Thus, for the selected $\boldsymbol{a}^*$, the optimization in the upper bound in (38) over $\boldsymbol{S} \succeq \boldsymbol{0}$ reduces to an optimization over $\boldsymbol{q}$, as $\boldsymbol{S} = P\boldsymbol{q}\boldsymbol{q}^T$

$$\max_{\boldsymbol{S}\succeq\boldsymbol{0}, \mathrm{tr}(\boldsymbol{S})\leq P} U(\boldsymbol{S}, \boldsymbol{a}^*) = \max_{\boldsymbol{q}} \frac{1}{2}\log\frac{\boldsymbol{q}^T(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H} + P\theta(\alpha^*)\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T)\boldsymbol{q}}{\boldsymbol{q}^T(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)\boldsymbol{q}} \tag{66}$$

where (66) is again in the Rayleigh quotient [26, p. 176] form, and the solution to this optimization problem is the largest eigenvalue of the matrix

$$(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1/2}(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H} + P\theta(\alpha^*)\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T) \times (\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1/2} \tag{67}$$

which is the largest eigenvalue of the matrix

$$(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H} + P\theta(\alpha^*)\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T) \tag{68}$$

since the two matrices are related by a similarity transformation. The eigenvalues of the matrix in (68) are given in the following lemma.

Lemma 2: The eigenvalues of the matrix in (68) are λ_1 and 1.

Proof: Note that

$$(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H} + P\theta(\alpha^*)\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T)\boldsymbol{q}_a = (\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}\left(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H}\right)\boldsymbol{q}_a \tag{69}$$

$$= \lambda_1\boldsymbol{q}_a \tag{70}$$

where (70) follows from (33).

Let us define vector $\boldsymbol{q}_1$ as

$$\boldsymbol{q}_1 = -\theta(\alpha^*)(\boldsymbol{H}^T\boldsymbol{H} - \boldsymbol{g}\boldsymbol{g}^T)^{-1}\boldsymbol{q}_a^\perp. \tag{71}$$

Note that

$$\boldsymbol{q}_1^T\boldsymbol{q}_a^\perp = -\theta(\alpha^*)(\boldsymbol{q}_a^\perp)^T(\boldsymbol{H}^T\boldsymbol{H} - \boldsymbol{g}\boldsymbol{g}^T)^{-1}\boldsymbol{q}_a^\perp = 1 \tag{72}$$

where the last equality follows from (64). Also, (71) implies that

$$\boldsymbol{H}^T\boldsymbol{H}\boldsymbol{q}_1 = \boldsymbol{g}\boldsymbol{g}^T\boldsymbol{q}_1 - \theta(\alpha^*)\boldsymbol{q}_a^\perp. \tag{73}$$

Then, we have

$$(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}(\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H} + P\theta(\alpha^*)\boldsymbol{q}_a^\perp(\boldsymbol{q}_a^\perp)^T)\boldsymbol{q}_1$$

$$= (\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}((\boldsymbol{I} + P\boldsymbol{H}^T\boldsymbol{H})\boldsymbol{q}_1 + P\theta(\alpha^*)\boldsymbol{q}_a^\perp) \tag{74}$$

$$= (\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1} \times (\boldsymbol{q}_1 + P\boldsymbol{g}\boldsymbol{g}^T\boldsymbol{q}_1 - P\theta(\alpha^*)\boldsymbol{q}_a^\perp + P\theta(\alpha^*)\boldsymbol{q}_a^\perp) \tag{75}$$

$$= (\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)^{-1}(\boldsymbol{I} + P\boldsymbol{g}\boldsymbol{g}^T)\boldsymbol{q}_1 \tag{76}$$

$$= \boldsymbol{q}_1 \tag{77}$$

where (74) follows from (72), and (75) follows from (73). From (70) and (77), we see that the two eigenvectors of the matrix in (68) are $\boldsymbol{q}_a$ and $\boldsymbol{q}_1$, and the corresponding eigenvalues are λ_1 and 1. □

Lemma 2 indicates that the eigenvalues of the matrix in (68), and also the eigenvalues of the matrix in (67), are λ_1 and 1. Since $\lambda_1 > 1$, as shown in (37), the resulting maximum value in (66) is $\frac{1}{2}\log\lambda_1$. Hence, the upper bound on the secrecy capacity, i.e., $\max_{\boldsymbol{S}\succeq\boldsymbol{0}, \mathrm{tr}(\boldsymbol{S})\leq P} U(\boldsymbol{S}, \boldsymbol{a}^*)$, is $\frac{1}{2}\log\lambda_1$, which is equal to the lower bound on the secrecy capacity shown in (35).

V. Conclusion

We determined the secrecy capacity of the 2-2-1 Gaussian MIMO wiretap channel, by solving for the optimum joint distribution for the auxiliary random variable and the channel input in the Csiszár–Körner formula. First, we proposed a lower bound on the secrecy capacity by evaluating the Csiszár–Körner formula for a specific selection of the auxiliary random variable and the channel input. Our achievable scheme is based on Gaussian

signalling and no pre-processing of information. Even for this achievable scheme, which is completely characterized by the transmit covariance matrix $\boldsymbol{S}$, a closed-form solution for the secrecy rate does not exist. However, in our 2-2-1 case, we have shown that the optimal transmission scheme is unit-rank, i.e., beam-forming is optimal.

We showed the optimality of the proposed achievable scheme by constructing a tight upper bound that meets it. The upper bound is developed by considering the secrecy capacity of a channel where the eavesdropper's signal is given to the legitimate receiver. Even though this upper bound is well-defined for a general MIMO wiretap channel, explicit evaluation and tightening of this upper bound has been possible by restricting ourselves to the 2-2-1 case. As in the lower bound, and by selecting a certain correlation structure for the additive noises, we have shown that beam-forming is optimal for the upper bound as well. Furthermore, we have shown that the optimal beam-forming directions in the lower and upper bounds are the same. Finally, we have shown that the two bounds meet yielding the secrecy capacity.

Our derivation is specific to the 2-2-1 case and we have not been able to show that these lower and upper bounds meet in the general MIMO channel. This is because the unit-rank (beam-forming) property of the optimum transmit matrices is essential in our derivations, while beam-forming is not likely to be the optimal strategy when the number of transmit and receive antennas is more than two. Shortly after the submission of this work, the secrecy capacity of the general Gaussian MIMO wiretap channel has been found by exploring the properties of the saddle point of the upper bound [27], [28]. Later, an alternataive approach to the general problem is developed using the channel enhancement technique [29].

APPENDIX

Proof of Theorem 1: A proof of similar results is presented for the case of m-1-n system, $m, n \geq 1$, in [10, Lemmas 1 and 2]. Our proof utilizes [10, Lemma 1], which generalizes to the case of multiple antennas at the legitimate receiver easily, and extends [10, Lemma 2] to the case where there are two antennas at the legitimate receiver.

An upper bound on the secrecy capacity of the wiretap channel described in (1) and (2) is [10, Lemma 1]

$$\max_{p(\boldsymbol{x}):E[\boldsymbol{x}^T\boldsymbol{x}]\leq P} I(\boldsymbol{x};\boldsymbol{y}\,|\,z). \tag{78}$$

Hence, we have

$$I(\boldsymbol{x};\boldsymbol{y}\,|\,z) = I(\boldsymbol{x};\boldsymbol{y},z) - I(\boldsymbol{x};z). \tag{79}$$

Intuitively, the upper bound is obtained by considering the secrecy capacity of a new channel where the legitimate receiver also has access to the eavesdropper's signal. Since the legitimate user is more capable in the new channel, the secrecy capacity of the new channel will serve as an upper bound on the secrecy capacity of the original channel. The new channel is degraded, and therefore the secrecy capacity formula is (79), obtained by setting $\boldsymbol{u} = \boldsymbol{x}$ as shown in [3].

In evaluating the right-hand side of (78), we introduce correlation between $\boldsymbol{n}_y$ and n_z, i.e., let us define $\boldsymbol{a}$ to be

$$\boldsymbol{a} = E[\boldsymbol{n}_y n_z]. \tag{80}$$

We note that $\boldsymbol{a}$ thus defined has to satisfy $||\boldsymbol{a}|| \leq 1$. To avoid irregular cases, we will only consider $\boldsymbol{a}$ such that $||\boldsymbol{a}|| < 1$. We also note that $\boldsymbol{a}$ does not affect the secrecy capacity of the original channel, but it affects the upper bound in (78). Thus, (78) remains an upper bound for any $\boldsymbol{a}$ with $||\boldsymbol{a}|| < 1$.

We evaluate $I(\boldsymbol{x};\boldsymbol{y}\,|\,z)$ as follows:

$$I(\boldsymbol{x};\boldsymbol{y}\,|\,z) = h(\boldsymbol{y}\,|\,z) - h(\boldsymbol{y}\,|\,z,\boldsymbol{x}) \tag{81}$$

$$= h(\boldsymbol{y}\,|\,z) - h(\boldsymbol{n}_y\,|\,n_z). \tag{82}$$

Due to the Gaussianity of the noise

$$h(\boldsymbol{n}_y\,|\,n_z) = h(\boldsymbol{n}_y, n_z) - h(n_z) = \frac{1}{2}\log(2\pi e)^2|\boldsymbol{N}| \tag{83}$$

where $\boldsymbol{N}$ is defined as in (40). Let us define $\boldsymbol{S}$ as

$$\boldsymbol{S} = E[\boldsymbol{x}\boldsymbol{x}^T] \tag{84}$$

then

$$E[\boldsymbol{y}z] = E[(\boldsymbol{H}\boldsymbol{x} + \boldsymbol{n}_y)(\boldsymbol{x}^T\boldsymbol{g} + n_z)] = \boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a} \tag{85}$$

$$E[z^2] = 1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g} \tag{86}$$

$$E[\boldsymbol{y}\boldsymbol{y}^T] = \boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T. \tag{87}$$

The linear minimum mean-squared error (LMMSE) estimator of $\boldsymbol{y}$ using z is

$$\hat{\boldsymbol{y}} = \frac{\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a}}{1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}} z \tag{88}$$

and the resulting covariance matrix of the estimation error is

$$\boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T - \frac{1}{1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}}(\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a})(\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a})^T. \tag{89}$$

Hence

$$h(\boldsymbol{y}\,|\,z) = h\left(\boldsymbol{y} - \frac{\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a}}{1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}} z \,\middle|\, z\right) \tag{90}$$

$$\leq h\left(\boldsymbol{y} - \frac{\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a}}{1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}} z\right) \tag{91}$$

$$\leq \frac{1}{2}\log(2\pi e)^2 \left| \boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T - \frac{1}{1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}} (\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a}) (\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a})^T \right|. \tag{92}$$

Therefore

$$I(\boldsymbol{x};\boldsymbol{y}|z) \leq \frac{1}{2}\log \frac{\left| \boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T - \frac{1}{1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g}}(\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a})(\boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a})^T \right|}{|\boldsymbol{N}|} \tag{93}$$

$$= \frac{1}{2}\log \frac{\left| \begin{bmatrix} \boldsymbol{I} + \boldsymbol{H}\boldsymbol{S}\boldsymbol{H}^T & \boldsymbol{H}\boldsymbol{S}\boldsymbol{g} + \boldsymbol{a} \\ \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{H}^T + \boldsymbol{a}^T & 1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g} \end{bmatrix} \right|}{(1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})|\boldsymbol{N}|} \tag{94}$$

$$= \frac{1}{2}\log \frac{|\boldsymbol{N} + \bar{\boldsymbol{H}}\boldsymbol{S}\bar{\boldsymbol{H}}^T|}{(1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})|\boldsymbol{N}|} \tag{95}$$

$$= \frac{1}{2}\log \frac{\left| \boldsymbol{I} + \boldsymbol{N}^{-1}\bar{\boldsymbol{H}}\boldsymbol{S}\bar{\boldsymbol{H}}^T \right|}{(1 + \boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})} \tag{96}$$

where $\bar{\boldsymbol{H}}$ is defined as in (41). Thus, we have

$$\max_{p(\boldsymbol{x}):E[\boldsymbol{x}^T\boldsymbol{x}]\le P} I(\boldsymbol{x};\boldsymbol{y}|z) \le \max_{\boldsymbol{S}\succeq\boldsymbol{0},\mathrm{tr}(\boldsymbol{S})\le P} \frac{1}{2}\log\frac{|\boldsymbol{I}+\boldsymbol{N}^{-1}\bar{\boldsymbol{H}}\boldsymbol{S}\bar{\boldsymbol{H}}^T|}{(1+\boldsymbol{g}^T\boldsymbol{S}\boldsymbol{g})}. \quad (97)$$

Therefore, an upper bound on the secrecy capacity of the wiretap channel described in (1) and (2) is

$$\max_{\boldsymbol{S}\succeq\boldsymbol{0},\mathrm{tr}(\boldsymbol{S})\le P} U(\boldsymbol{S},\boldsymbol{a}) \quad (98)$$

for any $\boldsymbol{a}$ with $||\boldsymbol{a}|| < 1$, with $U(\boldsymbol{S},\boldsymbol{a})$ defined in (39). □

REFERENCES

[1] A. D. Wyner, "The wiretap channel," *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 2–10, Oct. 1975.

[2] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wiretap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.

[3] I. Csiszár and J. Körner, "Broadcast channels with confidential messages," *IEEE Trans. Inf. Theory* , vol. IT-24, no. 3, pp. 339–348, May 1978.

[4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[5] R. Negi and S. Goel, "Secret communication using artificial noise," in *Proc. IEEE Vehicular Technology Conf.*, Toulouse, France, May 2006.

[6] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," in *Proc. IEEE Int. Symp. Information Theory*, Adelaide, Australia, Sep. 2005, pp. 2152–2155.

[7] S. Shafiee and S. Ulukus, "Achievable rates in Gaussian MISO channels with secrecy constraints," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 2466–2470.

[8] Z. Li, W. Trappe, and R. D. Yates, "Secret communication via multi-antenna transmission," in *Proc. 41st Conf. Information Science and Systems*, Baltimore, MD, Mar. 2007.

[9] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 2471–2475.

[10] A. Khisti and G. Wornell, "Secure transmission with multiple antennas: The MISOME wiretap channel," *IEEE Trans. Inf. Theory*, submitted for publication.

[11] Y. Liang and H. V. Poor, "Generalized multiple access channels with confidential messages," *IEEE Trans. Inf. Theory*, submitted for publication.

[12] R. Liu, I. Maric, R. D. Yates, and P. Spasojevic, "The discrete memoryless multiple access channel with confidential messages," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 957–961.

[13] E. Tekin and A. Yener, "Achievable rates for the general Gaussian multiple access wiretap channel with collective secrecy," in *Proc. 44th Ann. Allerton Conf. Commununications, Control, and Computing*, Monticello, IL, Sep. 2006.

[14] E. Tekin and A. Yener, "The Gaussian multiple access wiretap channel," *IEEE Trans. Inf. Theory*, submitted for publication.

[15] X. Tang, R. Liu, P. Spasojevic, and H. V. Poor, "Multiple access channels with generalized feedback and confidential messages," in *Proc. IEEE Inf. Theory Workshop on Frontiers in Coding Theory*, Lake Tahoe, CA, Sep. 2007.

[16] R. Liu and H. V. Poor, "Multiple antenna secure broadcast over wireless networks," in *Proc. 1st Int. Workshop on Information Theory Sens. Netw.*, Santa Fe, NM, Jun. 2007.

[17] L. Lai and H. El Gamal, "The relay-eavesdropper channel: Cooperation for secrecy," *IEEE Trans. Inf. Theory*, submitted for publication.

[18] Y. Oohama, "Relay channels with confidential messages," *IEEE Trans. Inf. Theory, Special Issue on Information Theoretic Security*, submitted for publication.

[19] 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, Special Issue Inf. Theoretic Security*, submitted for publication.

[20] E. Tekin and A. Yener, "The general Gaussian multiple access and two-way wiretap channels: Achievable rates and cooperative jamming," *IEEE Trans. Inf. Theory*, submitted for publication.

[21] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 356–360.

[22] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory, Special Issue on Inf. Theoretic Security*, submitted for publication.

[23] Z. Li, R. D. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," in *Proc. 44th Annual Allerton Conf. Communications, Control, and Computing*, Monticello, IL, Sep. 2006.

[24] P. K. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans. Inf. Theory*, submitted for publication.

[25] Z. Li, R. D. Yates, and W. Trappe, "Secure communication with a fading eavesdropper channel," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 1296–1300.

[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1999.

[27] A. Khisti and G. Wornell, "The MIMOME channel," in *Proc. 45th Ann. Allerton Conf. Communications, Control, and Computing*, Sep. 2007.

[28] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," in *Proc. 45th Ann. Allerton Conf. Communications, Control, and Computing*, Sep. 2007.

[29] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," available online: arXiv:cs.IT/0710.4105.

Shabnam Shafiee received the B.S. degree in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1996. She received the M.S. and Ph.D. degrees in electrical engineering from the University of Maryland, College Park, in 2006 and 2008, respectively.

She is now a Biomedical Engineer with Perinatronics Medical Systems, Millersville, MD.

Nan Liu (M'07) received the B.Eng. degree in electrical engineering from Beijing University of Posts and Telecommunications, Beijing, P. R. China, in 2001, and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, in 2007.

From 2007 to 2008, she was a postdoctoral scholar in the Wireless Systems Lab, Department of Electrical Engineering, Stanford University, Stanford, CA. In 2009, she became a faculty member of the School of Information Science and Engineering, Southeast University, Nanjing, China. Her research interests are in network information theory for wireless networks and wireless communication theory.

Sennur Ulukus (S'90–M'98) received the B.S. and M.S. degrees in electrical and electronics engineering from Bilkent University, Ankara, Turkey, in 1991 and 1993, respectively, and the Ph.D. degree in electrical and computer engineering from Rutgers University, New Brunswick, NJ, in 1998.

During her Ph.D. studies, she was with the Wireless Information Network Laboratory (WINLAB), Rutgers University. From 1998 to 2001, she was a Senior Technical Staff Member with AT&T Labs-Research, NJ. In 2001, she joined the University of Maryland at College Park, where she is currently an Associate Professor with the Department of Electrical and Computer Engineering, with a joint appointment at the Institute for Systems Research (ISR). Her research interests are in wireless communication theory and networking, network information theory for wireless networks, signal processing for wireless communications, and security for multiuser wireless communications.

Dr. Ulukus is a recipient of the 2005 NSF CAREER Award, and a corecipient of the 2003 IEEE Marconi Prize Paper Award in Wireless Communications. She serves or has served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY since 2007, as an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS between 2003–2007, as a Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS during 2006–2008, as the Co-Chair of the Communication Theory Symposium at the 2007 IEEE Global Telecommunications Conference, as the Co-Chair of the Medium Access Control (MAC) Track at the 2008 IEEE Wireless Communications and Networking Conference, and as the Secretary of the IEEE Communication Theory Technical Committee (CTTC) since 2007.