Gaussian MIMO Multi-Receiver Wiretap Channel

Ersen Ekrem Sennur Ulukus
Department of Electrical and Computer Engineering
University of Maryland, College Park, MD 20742
ersen@umd.edu *ulukus@umd.edu*

***Abstract*— We consider the Gaussian multiple-input multiple-output (MIMO) multi-receiver wiretap channel, and derive the secrecy capacity region of this channel for the most general case. We first prove the secrecy capacity region of the degraded MIMO channel, in which all receivers have the same number of antennas, and the noise covariance matrices exhibit a positive semi-definite order. We then generalize this result to the aligned case, in which all receivers have the same number of antennas, however there is no order among the noise covariance matrices. We accomplish this task by using the channel enhancement technique. Finally, we find the secrecy capacity region of the general MIMO channel by using some limiting arguments on the secrecy capacity region of the aligned MIMO channel. We show that a variant of dirty-paper coding with Gaussian signals is optimal.**

I. Introduction

Recently, information theoretic secrecy has gathered a renewed interest, and multiuser extensions of the wiretap channel [1], [2] have been widely investigated. One natural extension of the wiretap channel to the multiuser setting is the case of *secure broadcasting*, where there is one transmitter which wants to have confidential communication with many legitimate users in the presence of an external eavesdropper. Hereafter, we call this channel model the *multi-receiver wiretap channel*. This channel model has been studied in [3]–[6]. In particular, in [4]–[6], the degraded multi-receiver wiretap channel is considered, and its secrecy capacity region is derived for an arbitrary number of users in [4], [5], and for two users in [6]. The Gaussian scalar multi-receiver wiretap channel is an instance of the degraded multi-receiver wiretap channel.

The secrecy capacity region of the Gaussian scalar multi-receiver wiretap channel is found in [7], [8]. Furthermore, in [7], [8], it is shown that existing converse techniques for the Gaussian scalar broadcast channel, i.e., the ones in [9], [10], cannot be extended in a straightforward manner to provide a converse proof for the Gaussian scalar multi-receiver wiretap channel. The main reason for this is the insufficiency of the entropy-power inequality [11] in resolving the ambiguity regarding the auxiliary random variables. Two converse proofs are provided in [7], [8]. The first one uses the relationship between the mutual information and the minimum-mean-square-error (MMSE) along with the properties of the MMSE. This relationship is also used in [12] to prove the secrecy capacity of the Gaussian MIMO wiretap channel. The second one uses the relationship between the differential entropy and the Fisher information along with the properties of the Fisher information. This reveals that either the Fisher information matrix or the MMSE matrix should play an important role in the converse proof of the MIMO case.

Keeping this motivation in mind, we consider the Gaussian MIMO multi-receiver wiretap channel here. Instead of directly tackling the most general case, we first consider two sub-classes of MIMO channels. In the first sub-class, all receivers have the same number of antennas and the noise covariance matrices exhibit a positive semi-definite order, which implies the degradedness of these channels. Hereafter, we call this channel model the *degraded Gaussian MIMO multi-receiver wiretap channel*. In the second sub-class, although all receivers still have the same number of antennas as in the degraded case, the noise covariance matrices do not have to satisfy any positive semi-definite order. Hereafter, we call this channel model the *aligned Gaussian MIMO multi-receiver wiretap channel*. Our approach will be to first find the secrecy capacity region of the degraded case, then to generalize this result to the aligned case by using the *channel enhancement* technique [13]. Once we obtain the secrecy capacity region of the aligned case, we use this result to find the secrecy capacity region of the most general case by some limiting arguments as in [13], [14]. Thus, the main contribution and the novelty of our work is the way we prove the secrecy capacity region of the degraded Gaussian MIMO multi-receiver wiretap channel.

To clarify our contributions, it is useful to note the similarity of the proof steps that we follow here with those in [13], where the capacity region of the Gaussian MIMO broadcast channel was established. In [13] also, the authors considered the degraded, the aligned and the general cases successively. Although, both [13] and this paper have these same proof steps, there are differences between how and why these steps are taken. In [13], the main difficulty in obtaining the capacity region of the Gaussian MIMO broadcast channel was to extend Bergmans' converse for the scalar case to the degraded vector channel. This difficulty was overcome in [13] by the invention of the *channel enhancement* technique. However, as discussed earlier, as shown in [7], [8], Bergmans' converse cannot be extended to our secrecy context, even for the degraded scalar case. Thus, we need a new proof technique which we construct by using the Fisher information matrix and the generalized de Bruin identity [15]. After we obtain the secrecy capacity region of the degraded MIMO channel, we adapt the channel enhancement technique to our setting to find the secrecy capacity region of the aligned MIMO channel. The difference

This work was supported by NSF Grants CCF 04-47613, CCF 05-14846, CNS 07-16311 and CCF 07-29127.

of the way channel enhancement is used here as compared to the one in [13] comes from the presence of an eavesdropper, and its difference from the one in [14] is due to the presence of many legitimate users. After we find the secrecy capacity region of the aligned MIMO channel, we use the limiting arguments that appeared in [13], [14] to prove the secrecy capacity region of the general MIMO channel.

II. Multi-Receiver Wiretap Channels

The multi-receiver wiretap channel consists of one transmitter, K legitimate users and an eavesdropper. The transmitter sends a confidential message to each user and all messages are to be kept secret from the eavesdropper. The channel is memoryless with a transition probability $p(y_1, y_2, \ldots, y_K, z|x)$. A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code consists of K message sets, $\mathcal{W}_k = \{1, \ldots, 2^{nR_k}\}$, $k = 1, \ldots, K$, an encoder $f : \mathcal{W}_1 \times \ldots \times \mathcal{W}_K \rightarrow \mathcal{X}^n$, K decoders, $g_k : \mathcal{Y}_k \rightarrow \mathcal{W}_k$, $k = 1, \ldots, K$. A rate tuple $(R_1, \ldots, R_K)$ is said to be achievable if there exists a code with vanishingly small probability of error and

$$\lim_{n\rightarrow\infty} \frac{1}{n} H(\mathcal{S}(W)|Z^n) \geq \sum_{k\in\mathcal{S}(W)} R_k, \quad \forall\ \mathcal{S}(W) \qquad (1)$$

where $\mathcal{S}(W)$ denotes any subset of $\{W_1, \ldots, W_K\}$. Hence, we consider only perfect secrecy rates. The secrecy capacity region is defined as the closure of all achievable rate tuples.

The degraded multi-receiver wiretap channel exhibits the following Markov chain

$$X \rightarrow Y_1 \rightarrow \ldots \rightarrow Y_K \rightarrow Z \qquad (2)$$

whose capacity region was established in [4], [5] for an arbitrary number of users and in [6] for two users.

Theorem 1 *The secrecy capacity region of the degraded multi-receiver wiretap channel is given by the union of the rate tuples* $(R_1, \ldots, R_K)$ *satisfying*

$$R_k \leq I(U_k; Y_k|U_{k+1}, Z), \quad k = 1, \ldots, K \qquad (3)$$

where $U_1 = X, U_{K+1} = \phi$, *and the union is over all probability distributions of the form*

$$p(u_K)p(u_{K-1}|u_K)\ldots p(u_2|u_3)p(x|u_2) \qquad (4)$$

Since the channel is degraded, i.e., we have the Markov chain in (2), the capacity expressions in (3) are equivalent to

$$R_k \leq I(U_k; Y_k|U_{k+1}) - I(U_k; Z|U_{k+1}), \quad k = 1, \ldots, K \quad (5)$$

III. Gaussian MIMO Multi-receiver Wiretap Channel

A. Degraded Case

A degraded Gaussian MIMO multi-receiver wiretap channel is defined by

$$\mathbf{Y}_k = \mathbf{X} + \mathbf{N}_k, \quad k = 1, \ldots, K \qquad (6)$$

$$\mathbf{Z} = \mathbf{X} + \mathbf{N}_Z \qquad (7)$$

where the input is subject to a covariance constraint

$$E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S} \qquad (8)$$

where $\mathbf{S} \succ \mathbf{0}$, and the noise covariance matrices of the Gaussian random vectors $\{\mathbf{N}_k\}_{k=1}^K, \mathbf{N}_Z$, i.e., $\{\boldsymbol{\Sigma}_k\}_{k=1}^K, \boldsymbol{\Sigma}_Z$, satisfy the following semi-definite order

$$\mathbf{0} \prec \boldsymbol{\Sigma}_1 \preceq \boldsymbol{\Sigma}_2 \preceq \ldots \preceq \boldsymbol{\Sigma}_K \preceq \boldsymbol{\Sigma}_Z \qquad (9)$$

B. Aligned Case

An aligned Gaussian MIMO multi-receiver wiretap channel is defined by (6)-(7), and the input is subject to the same covariance constraint as in (8). The only assumption about the noise covariance matrices is that they are strictly positive definite, however they do not need to satisfy any positive semi-definite order as opposed to the degraded case.

C. General Case

The general form of the Gaussian MIMO multi-receiver wiretap channel is given by

$$\mathbf{Y}_k = \mathbf{H}_k\mathbf{X} + \mathbf{N}_k, \quad k = 1, \ldots, K \qquad (10)$$

$$\mathbf{Z} = \mathbf{H}_Z\mathbf{X} + \mathbf{N}_Z \qquad (11)$$

where the input $\mathbf{X}$, which is a $t \times 1$ column vector, is subject to the same covariance constraint as in (8). The kth user's observation $\mathbf{Y}_k$ is a column vector of size $r_k \times 1,\ k = 1, \ldots, K$. The eavesdropper's observation $\mathbf{Z}$ is of size $r_Z \times 1$. The covariance matrices of the Gaussian random vectors $\{\mathbf{N}_k\}_{k=1}^K, \mathbf{N}_Z$, i.e., $\{\boldsymbol{\Sigma}_k\}_{k=1}^K, \boldsymbol{\Sigma}_Z$, are again assumed to be strictly positive definite. The channel gain matrices $\{\mathbf{H}_k\}_{k=1}^K, \mathbf{H}_Z$ are of sizes $\{r_k \times t\}_{k=1}^K, r_Z \times t$, respectively, and they are known to the transmitter, all legitimate users and the eavesdropper.

IV. Main Results

A. Degraded Case

The order in (9) is equivalent to the following Markov chain

$$\mathbf{X} \rightarrow \mathbf{Y}_1 \rightarrow \ldots \rightarrow \mathbf{Y}_K \rightarrow \mathbf{Z} \qquad (12)$$

because of the fact that the capacity-equivocation rate region of a multi-receiver wiretap channel depends only on the conditional marginal distributions of the channel, but not on the entire joint distribution of the channel. Thus, Theorem 1 gives the secrecy capacity region for the degraded case. Later, we will show that Gaussian selections for the auxiliary random variables and the channel input in Theorem 1 is optimal.

Theorem 2 *The secrecy capacity region of the degraded Gaussian MIMO multi-receiver wiretap channel is given by the union of the rate tuples* $(R_1, \ldots, R_K)$ *satisfying*

$$R_k \leq \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i + \boldsymbol{\Sigma}_k\right|}{\left|\sum_{i=1}^{k-1}\mathbf{K}_i + \boldsymbol{\Sigma}_k\right|} - \frac{1}{2}\log\frac{\left|\sum_{i=1}^{k}\mathbf{K}_i + \boldsymbol{\Sigma}_Z\right|}{\left|\sum_{i=1}^{k-1}\mathbf{K}_i + \boldsymbol{\Sigma}_Z\right|}, \quad k = 1, \ldots, K \qquad (13)$$

where the union is over all positive semi-definite matrices $\{\mathbf{K}_i\}_{i=1}^K$ *that satisfy* $\sum_{i=1}^K \mathbf{K}_i = \mathbf{S}$.

The proof of this theorem for $K = 2$ is given in Section V. The proof for arbitrary K can be found in [8].

B. Aligned and General Cases

In both aligned and general cases, there is no single-letter characterization for the secrecy capacity region as opposed to the degraded case. Nevertheless, we are able to establish the secrecy capacity region of both cases. We show that dirty-paper coding with stochastic encoding is optimal in both cases.

Given the covariance matrices $\{\mathbf{K}_k\}_{k=1}^K$ such that $\sum_{k=1}^K \mathbf{K}_k \preceq \mathbf{S}$, we define the following rates

$$\begin{aligned} & R_k^{\text{DPC}}\left(\pi, \{\mathbf{K}_i\}_{i=1}^K\right) \\ & \quad = \frac{1}{2}\log \frac{\left|\mathbf{H}_{\pi(k)}\left(\sum_{i=1}^k \mathbf{K}_{\pi(i)}\right)\mathbf{H}_{\pi(k)}^\top + \mathbf{\Sigma}_{\pi(k)}\right|}{\left|\mathbf{H}_{\pi(k)}\left(\sum_{i=1}^{k-1} \mathbf{K}_{\pi(i)}\right)\mathbf{H}_{\pi(k)}^\top + \mathbf{\Sigma}_{\pi(k)}\right|} \\ & \qquad - \frac{1}{2}\log \frac{\left|\mathbf{H}_Z\left(\sum_{i=1}^k \mathbf{K}_{\pi(i)}\right)\mathbf{H}_Z^\top + \mathbf{\Sigma}_Z\right|}{\left|\mathbf{H}_Z\left(\sum_{i=1}^{k-1} \mathbf{K}_{\pi(i)}\right)\mathbf{H}_Z^\top + \mathbf{\Sigma}_Z\right|} \end{aligned} \tag{14}$$

for $k = 1, \ldots, K$, where $\pi(\cdot)$ is a one-to-one permutation on $\{1, \ldots, K\}$. We also note that the subscript of $R_k^{\text{DPC}}\left(\pi, \{\mathbf{K}_i\}_{i=1}^K\right)$ does not denote the kth user, instead it denotes the $(K-k+1)$th user in line to be encoded. Rather, the secrecy rate of the kth user is given by

$$R_k = R_{\pi^{-1}(k)}^{\text{DPC}}\left(\pi, \{\mathbf{K}_i\}_{i=1}^K\right) \tag{15}$$

when dirty-paper coding with stochastic encoding is used with an encoding order of π.

We define the following region:

$$\begin{aligned} & \mathcal{R}^{\text{DPC}}(\pi, \mathbf{S}) \\ & = \bigcup \left\{(R_1, \ldots, R_K) : R_k = R_{\pi^{-1}(k)}^{\text{DPC}}\left(\pi, \{\mathbf{K}_i\}_{i=1}^K\right)\right\} \end{aligned} \tag{16}$$

where the union is over all positive semi-definite matrices $\{\mathbf{K}_i\}_{i=1}^K$ that satisfy $\sum_{i=1}^K \mathbf{K}_i \preceq \mathbf{S}$.

The secrecy capacity region of the aligned and general Gaussian MIMO broadcast channels is given as follows.

Theorem 3 *The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel is given by the convex closure of the following union*

$$\bigcup_{\pi \in \Pi} \mathcal{R}^{\text{DPC}}(\pi, \mathbf{S}) \tag{17}$$

where Π is the set of all possible one-to-one permutations on $\{1, \ldots, K\}$.

Achievability of the region in Theorem 3 can be shown by using dirty-paper coding in conjunction with stochastic encoding [8]. We prove the converse statement in two steps. We first provide the converse proof for the aligned case, then for the general case. For the converse proof of the aligned case, we use the channel enhancement technique [13]. Basically, for each point on the boundary of the secrecy capacity region of any aligned channel, there exists a degraded channel such that the secrecy capacity region of this degraded channel includes the secrecy capacity region of the aligned channel, and both regions intersect at this specific point. We then complete the converse proof since we know the secrecy capacity region of the degraded channel due to Theorem 2. To provide the converse proof for the general case, we use some limiting arguments as in [13], [14]. The details of these proofs can be found in [8].

V. Proof of Theorem 2 for $K = 2$

For the upcoming proof, we need the results of some intermediate optimization problems. The first one is the so-called worst additive noise lemma [16], [17].

Lemma 1 *Let $\mathbf{N}$ be a Gaussian random vector with covariance matrix $\mathbf{\Sigma}$, and $\mathbf{K}_X$ be a positive semi-definite matrix. Consider the following optimization problem,*

$$\min_{p(\mathbf{x})} \quad I(\mathbf{N}; \mathbf{N} + \mathbf{X}) \qquad \text{s.t.} \quad \text{Cov}(\mathbf{X}) = \mathbf{K}_X \tag{18}$$

where $\mathbf{X}$ and $\mathbf{N}$ are independent. A Gaussian $\mathbf{X}$ is the minimizer of this optimization problem.

The second optimization problem that will be useful in the upcoming proof is given in the following theorem.

Theorem 4 *Let $\mathbf{U}, \mathbf{X}$ be arbitrarily correlated random vectors which are independent of $\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}_Z$, where $\mathbf{N}_1, \mathbf{N}_2, \mathbf{N}_Z$ are zero-mean Gaussian random vectors with covariance matrices $\mathbf{0} \prec \mathbf{\Sigma}_1 \preceq \mathbf{\Sigma}_2 \preceq \mathbf{\Sigma}_Z$, respectively. Moreover, assume that the second moment of $\mathbf{X}$ is constrained as*

$$E\left[\mathbf{X}\mathbf{X}^\top\right] \preceq \mathbf{S} \tag{19}$$

where $\mathbf{S}$ is a positive definite matrix. Then, for any admissible $(\mathbf{U}, \mathbf{X})$ pair, there exists a matrix $\mathbf{K}^*$ such that $\mathbf{0} \preceq \mathbf{K}^* \preceq \mathbf{S}$, and

$$h(\mathbf{X} + \mathbf{N}_Z|\mathbf{U}) - h(\mathbf{X} + \mathbf{N}_2|\mathbf{U}) = \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{20}$$

$$h(\mathbf{X} + \mathbf{N}_Z|\mathbf{U}) - h(\mathbf{X} + \mathbf{N}_1|\mathbf{U}) \geq \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_1|} \tag{21}$$

We are now ready to prove the secrecy capacity region of the two-user degraded MIMO Gaussian multi-receiver channel. We first consider R_2, and bound it using Theorem 1 as follows

$$R_2 \leq I(U_2; \mathbf{Y}_2) - I(U_2; \mathbf{Z}) \tag{22}$$

$$\begin{aligned} & = [I(\mathbf{X}; \mathbf{Y}_2) - I(\mathbf{X}; \mathbf{Z})] \\ & \quad - [I(\mathbf{X}; \mathbf{Y}_2|U_2) - I(\mathbf{X}; \mathbf{Z}|U_2)] \end{aligned} \tag{23}$$

where the equality is obtained by using the chain rule and the Markov chain $U_2 \to \mathbf{X} \to (\mathbf{Y}_2, \mathbf{Z})$. The expression in the first bracket of (23) is

$$I(\mathbf{X}; \mathbf{Y}_2) - I(\mathbf{X}; \mathbf{Z}) = h(\mathbf{Y}_2) - h(\mathbf{Z}) - \frac{1}{2}\log\frac{|\mathbf{\Sigma}_2|}{|\mathbf{\Sigma}_Z|} \tag{24}$$

We consider the difference of differential entropies in (24). To this end, consider the Gaussian random vector $\tilde{\mathbf{N}}_2$ with covariance matrix $\mathbf{\Sigma}_Z - \mathbf{\Sigma}_2$, which is chosen to be independent of $\mathbf{X}, \mathbf{N}_2$. Using the Markov chain in (12), we get

$$
\begin{aligned}
h(\mathbf{Y}_2) - h(\mathbf{Z}) &= h(\mathbf{Y}_2) - h(\mathbf{Y}_2 + \tilde{\mathbf{N}}_2) && (25)\\
&= -I(\tilde{\mathbf{N}}_2; \mathbf{Y}_2 + \tilde{\mathbf{N}}_2) && (26)\\
&\leq \max_{\mathbf{0} \preceq \mathbf{K} \preceq \mathbf{S}} \frac{1}{2} \log \frac{|\mathbf{K} + \mathbf{\Sigma}_2|}{|\mathbf{K} + \mathbf{\Sigma}_Z|} && (27)\\
&= \frac{1}{2} \log \frac{|\mathbf{S} + \mathbf{\Sigma}_2|}{|\mathbf{S} + \mathbf{\Sigma}_Z|} && (28)
\end{aligned}
$$

where (27) follows from Lemma 1 and (28) is a consequence of the fact that the objective function in (27) is monotonically increasing in $\mathbf{K}$ [13]. Plugging (28) into (24) yields

$$
I(\mathbf{X}; \mathbf{Y}_2) - I(\mathbf{X}; \mathbf{Z}) \leq \frac{1}{2} \log \frac{|\mathbf{S} + \mathbf{\Sigma}_2|}{|\mathbf{\Sigma}_2|} - \frac{1}{2} \log \frac{|\mathbf{S} + \mathbf{\Sigma}_Z|}{|\mathbf{\Sigma}_Z|} \tag{29}
$$

We now consider the expression in the second bracket of (23). For that purpose, we use Theorem 4. According to Theorem 4, for any admissible pair $(U_2, \mathbf{X})$, there exists a $\mathbf{K}^*$ such that

$$
h(\mathbf{Z}|U_2) - h(\mathbf{Y}_2|U_2) = \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{30}
$$

which is equivalent to

$$
\begin{aligned}
I(\mathbf{X}; \mathbf{Z}|U_2) - I(\mathbf{X}; \mathbf{Y}_2|U_2) = \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{\Sigma}_Z}&\\
- \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_2|}{|\mathbf{\Sigma}_2|}& \qquad (31)
\end{aligned}
$$

Thus, using (29) and (31) in (23), we get

$$
R_2 \leq \frac{1}{2} \log \frac{|\mathbf{S} + \mathbf{\Sigma}_2|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} - \frac{|\mathbf{S} + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_Z|} \tag{32}
$$

which is the desired bound on R_2 given in Theorem 2.

We now bound R_1. To this end, we first bound R_1 using Theorem 1

$$
\begin{aligned}
R_1 &\leq I(\mathbf{X}; \mathbf{Y}_1|U_2) - I(\mathbf{X}; \mathbf{Z}|U_2) && (33)\\
&= h(\mathbf{Y}_1|U_2) - h(\mathbf{Z}|U_2) - \frac{1}{2} \log \frac{|\mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_Z|} && (34)
\end{aligned}
$$

To bound the difference of conditional differential entropies in (34), we use Theorem 4. Theorem 4 states that for any admissible pair $(U_2, \mathbf{X})$, there exists a matrix $\mathbf{K}^*$ such that it satisfies (30) and also

$$
h(\mathbf{Z}|U_2) - h(\mathbf{Y}_1|U_2) \geq \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_1|} \tag{35}
$$

Thus, using (35) in (34), we get

$$
R_1 \leq \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_1|}{|\mathbf{\Sigma}_1|} - \frac{1}{2} \log \frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{\Sigma}_Z|} \tag{36}
$$

which is the desired bound on R_1 given in Theorem 2, completing the converse proof for $K = 2$.

As we have seen, the main ingredient in the above proof was Theorem 4. Therefore, to complete the converse proof for the degraded channel for $K = 2$, from this point on, we need the proof of Theorem 4, which will be provided next.

VI. Proof of Theorem 4

Due to space limitations here, we provide the proof of Theorem 4 for the case $\mathbf{U} = \phi$; the proof for the general case can be found in [8]. The upcoming proof involves the use of the properties of Fisher information, and its connection to the differential entropy, which are provided next.

A. The Fisher Information Matrix

We start with the definition.

Definition 1 *Let $\mathbf{U}$ be a length-n random vector with differentiable density $f_U(\mathbf{u})$. The Fisher information matrix of $\mathbf{U}$, $\mathbf{J}(\mathbf{U})$, is defined as*

$$
\mathbf{J}(\mathbf{U}) = E\left[\boldsymbol{\rho}(\mathbf{U})\boldsymbol{\rho}(\mathbf{U})^\top\right] \tag{37}
$$

where $\boldsymbol{\rho}(\mathbf{u})$ is the score function which is given by

$$
\boldsymbol{\rho}(\mathbf{u}) = \nabla \log f_U(\mathbf{u}) = \left[\frac{\partial \log f_U(\mathbf{u})}{\partial u_1} \ \ldots \ \frac{\partial \log f_U(\mathbf{u})}{\partial u_n}\right]^\top \tag{38}
$$

The following lemma, which is due to [18], is instrumental in the upcoming proof.

Lemma 2 ([18]) *Let $\mathbf{U}$ be a random vector with differentiable density, and let $\mathbf{\Sigma}_U \succ \mathbf{0}$ be its covariance matrix. Then, we have*

$$
\mathbf{J}(\mathbf{U}) \succeq \mathbf{\Sigma}_U^{-1} \tag{39}
$$

which is satisfied with equality if $\mathbf{U}$ is Gaussian.

We will need the following lemma as well.

Lemma 3 ([8]) *Let $\mathbf{U}$ and $\mathbf{V}$ be independent random vectors with differentiable densities. Then, we have*

1) $\mathbf{J}(\mathbf{U} + \mathbf{V}) \preceq \mathbf{J}(\mathbf{U})$
2) $\mathbf{J}(\mathbf{U} + \mathbf{V}) \preceq \left[\mathbf{J}(\mathbf{U})^{-1} + \mathbf{J}(\mathbf{V})^{-1}\right]^{-1}$

The following lemma regarding the Fisher information matrix is also useful in the upcoming proof.

Lemma 4 ([8]) *Let $\mathbf{U}, \mathbf{V}_1, \mathbf{V}_2$ be random vectors such that $\mathbf{U}$ and $(\mathbf{V}_1, \mathbf{V}_2)$ are independent. Moreover, let $\mathbf{V}_1, \mathbf{V}_2$ be Gaussian random vectors with covariance matrices $\mathbf{0} \prec \mathbf{\Sigma}_1 \preceq \mathbf{\Sigma}_2$. Then, we have*

$$
\mathbf{J}(\mathbf{U} + \mathbf{V}_2)^{-1} - \mathbf{\Sigma}_2 \succeq \mathbf{J}(\mathbf{U} + \mathbf{V}_1)^{-1} - \mathbf{\Sigma}_1 \tag{40}
$$

We need the relationship between the Fisher information matrix and the differential entropy, which is due to [15].

Lemma 5 ([15]) *Let $\mathbf{X}$ and $\mathbf{N}$ be independent random vectors, where $\mathbf{N}$ is zero-mean Gaussian with covariance matrix $\mathbf{\Sigma}_N \succ \mathbf{0}$, and $\mathbf{X}$ has a finite second order moment. Then,*

$$
\nabla_{\mathbf{\Sigma}_N} h(\mathbf{X} + \mathbf{N}) = \frac{1}{2}\mathbf{J}(\mathbf{X} + \mathbf{N}) \tag{41}
$$

B. Proof of Theorem 4 for $\mathbf{U}=\phi$

To prove Theorem 4, we first consider the following expression

$$h(\mathbf{X}+\mathbf{N}_Z)-h(\mathbf{X}+\mathbf{N}_2) \tag{42}$$

which is bounded due to the covariance constraint on $\mathbf{X}$. In particular, we have

$$\frac{1}{2}\log\frac{|\mathbf{S}+\boldsymbol{\Sigma}_Z|}{|\mathbf{S}+\boldsymbol{\Sigma}_2|} \le h(\mathbf{X}+\mathbf{N}_Z)-h(\mathbf{X}+\mathbf{N}_2) \le \frac{1}{2}\log\frac{|\boldsymbol{\Sigma}_Z|}{|\boldsymbol{\Sigma}_2|} \tag{43}$$

where the lower bound can be obtained by using Lemma 1, and the upper bound can be derived by using the stability of Gaussian random vectors and the fact that conditioning cannot increase entropy [8]. Thus, we can fix the difference of the differential entropies in (43) to an α in this range, i.e.,

$$h(\mathbf{X}+\mathbf{N}_Z)-h(\mathbf{X}+\mathbf{N}_2)=\alpha \tag{44}$$

where $\alpha \in \left[\frac{1}{2}\log|\mathbf{S}+\boldsymbol{\Sigma}_Z|/|\mathbf{S}+\boldsymbol{\Sigma}_2|, \frac{1}{2}\log|\boldsymbol{\Sigma}_Z|/|\boldsymbol{\Sigma}_2|\right]$. We now would like to understand how the constraint in (44) affects the set of admissible random vectors. For that purpose, we use Lemma 5, and express this difference of entropies as an integral of the Fisher information matrix[1]

$$\alpha = h(\mathbf{X}+\mathbf{N}_Z)-h(\mathbf{X}+\mathbf{N}_2)=\frac{1}{2}\int_{\boldsymbol{\Sigma}_2}^{\boldsymbol{\Sigma}_Z}\mathbf{J}(\mathbf{X}+\mathbf{N})d\boldsymbol{\Sigma}_N \tag{45}$$

Using the stability of Gaussian random vectors, we can express $\mathbf{J}(\mathbf{X}+\mathbf{N})$ as

$$\mathbf{J}(\mathbf{X}+\mathbf{N})=\mathbf{J}(\mathbf{X}+\mathbf{N}_2+\tilde{\mathbf{N}}) \tag{46}$$

where $\tilde{\mathbf{N}}$ is a zero-mean Gaussian random vector with covariance matrix $\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_2\succeq\mathbf{0}$, and is independent of $\mathbf{N}_2$. Using the second part of Lemma 3 in (46), we get

$$\begin{aligned}\mathbf{J}(\mathbf{X}+\mathbf{N})&=\mathbf{J}(\mathbf{X}+\mathbf{N}_2+\tilde{\mathbf{N}}) &(47)\\ &\preceq\left[\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}+\mathbf{J}(\tilde{\mathbf{N}})^{-1}\right]^{-1} &(48)\\ &=\left[\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}+\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_2\right]^{-1} &(49)\end{aligned}$$

where we used the fact that $\mathbf{J}(\tilde{\mathbf{N}})=(\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_2)^{-1}$ which is a consequence of Lemma 2 by noting that $\tilde{\mathbf{N}}$ is Gaussian. We now bound the integral in (45) by using (49). For that purpose, we introduce the following lemma.

Lemma 6 ([8]) *Let $\mathbf{K}_1,\mathbf{K}_2$ be positive semi-definite matrices satisfying $\mathbf{0}\preceq\mathbf{K}_1\preceq\mathbf{K}_2$, and $\mathbf{f}(\mathbf{K})$ be a matrix-valued function such that $\mathbf{f}(\mathbf{K})\succeq\mathbf{0}$ for $\mathbf{K}_1\preceq\mathbf{K}\preceq\mathbf{K}_2$. Then, we have*

$$\int_{\mathbf{K}_1}^{\mathbf{K}_2}\mathbf{f}(\mathbf{K})d\mathbf{K}\ge 0 \tag{50}$$

In light of this lemma, using (49) in (45), we get

$$\begin{aligned}\alpha&\le\frac{1}{2}\int_{\boldsymbol{\Sigma}_2}^{\boldsymbol{\Sigma}_Z}\left[\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}+\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_2\right]^{-1}d\boldsymbol{\Sigma}_N &(51)\\ &=\frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}+\boldsymbol{\Sigma}_Z-\boldsymbol{\Sigma}_2|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}|} &(52)\end{aligned}$$

where we used the fact that $\nabla_{\boldsymbol{\Sigma}}\log|\boldsymbol{\Sigma}|=\boldsymbol{\Sigma}^{-\top}$ for $\boldsymbol{\Sigma}\succ\mathbf{0}$. We also note that the denominator in (52) is strictly positive, i.e., $|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}|>0$, because

$$\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}\succeq\mathbf{J}(\mathbf{N}_2)^{-1}=\boldsymbol{\Sigma}_2\succ\mathbf{0} \tag{53}$$

Following similar steps, we can also find a lower bound on α. Again, using the stability of Gaussian random vectors, we have

$$\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)=\mathbf{J}(\mathbf{X}+\mathbf{N}+\tilde{\mathbf{N}}) \tag{54}$$

where $\mathbf{N},\tilde{\mathbf{N}}$ are zero-mean Gaussian random vectors with covariance matrices $\boldsymbol{\Sigma}_N,\boldsymbol{\Sigma}_Z-\boldsymbol{\Sigma}_N$, respectively, $\boldsymbol{\Sigma}_2\preceq\boldsymbol{\Sigma}_N\preceq\boldsymbol{\Sigma}_Z$, and they are independent. Using the second part of Lemma 3 in (54) yields

$$\begin{aligned}\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)&=\mathbf{J}(\mathbf{X}+\mathbf{N}+\tilde{\mathbf{N}}) &(55)\\ &\preceq\left[\mathbf{J}(\mathbf{X}+\mathbf{N})^{-1}+\mathbf{J}(\tilde{\mathbf{N}})^{-1}\right]^{-1} &(56)\\ &=\left[\mathbf{J}(\mathbf{X}+\mathbf{N})^{-1}+\boldsymbol{\Sigma}_Z-\boldsymbol{\Sigma}_N\right]^{-1} &(57)\end{aligned}$$

where we used the fact that $\mathbf{J}(\tilde{\mathbf{N}})=(\boldsymbol{\Sigma}_Z-\boldsymbol{\Sigma}_N)^{-1}$ which follows from Lemma 2 due to the Gaussianity of $\tilde{\mathbf{N}}$. By noting the fact that if $\mathbf{A}\succeq\mathbf{B}\succ\mathbf{0}$, then $\mathbf{B}^{-1}\succeq\mathbf{A}^{-1}\succ\mathbf{0}$, (57) is equivalent to

$$\left[\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_Z\right]^{-1}\preceq\mathbf{J}(\mathbf{X}+\mathbf{N}) \tag{58}$$

Use of Lemma 6 and (58) in (45) yields

$$\begin{aligned}\alpha&\ge\int_{\boldsymbol{\Sigma}_2}^{\boldsymbol{\Sigma}_Z}\left[\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_N-\boldsymbol{\Sigma}_Z\right]^{-1}d\boldsymbol{\Sigma}_N &(59)\\ &=\frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_Z|} &(60)\end{aligned}$$

where we again used $\nabla_{\boldsymbol{\Sigma}}\log|\boldsymbol{\Sigma}|=\boldsymbol{\Sigma}^{-\top}$ for $\boldsymbol{\Sigma}\succ\mathbf{0}$. Here also, the denominator is strictly positive, i.e., $|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_Z|>0$, because

$$\begin{aligned}\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_Z&\succeq\mathbf{J}(\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_Z\\ &=\boldsymbol{\Sigma}_2\succ\mathbf{0} &(61)\end{aligned}$$

Combining the two bounds on α given in (52) and (60) yields

$$\begin{aligned}\frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}+\boldsymbol{\Sigma}_2-\boldsymbol{\Sigma}_Z|}&\le\alpha\\ &\le\frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}+\boldsymbol{\Sigma}_Z-\boldsymbol{\Sigma}_2|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}|} &(62)\end{aligned}$$

Next, we will need the following lemma.

Lemma 7 ([8]) *Consider the function*

$$r(t)=\frac{1}{2}\log\frac{|\mathbf{A}+\mathbf{B}+t\boldsymbol{\Delta}|}{|\mathbf{A}+t\boldsymbol{\Delta}|},\qquad 0\le t\le 1 \tag{63}$$

[1] $\int_{\boldsymbol{\Sigma}_2}^{\boldsymbol{\Sigma}_Z}\mathbf{J}(\cdot)d\boldsymbol{\Sigma}$ denotes the line integral of the vector-valued function $\mathbf{J}(\cdot)$ over any path from $\boldsymbol{\Sigma}_2$ to $\boldsymbol{\Sigma}_Z$. Since $\mathbf{J}(\cdot)$ is the gradient of a scalar field, this integration is path-free. This remark applies to all upcoming integrations of $\mathbf{J}(\cdot)$.

where $\mathbf{A}, \mathbf{B}, \mathbf{\Delta}$ are real, symmetric matrices, and $\mathbf{A} \succ \mathbf{0}$, $\mathbf{B} \succeq \mathbf{0}, \mathbf{\Delta} \succeq \mathbf{0}$. The function $r(t)$ is continuous and monotonically decreasing in t.

To investigate the implications of (62), let us select $\mathbf{A}, \mathbf{B}, \mathbf{\Delta}$ in $r(t)$ in Lemma 7 as follows

$$\mathbf{A} = \mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1} \tag{64}$$

$$\mathbf{B} = \mathbf{\Sigma}_Z - \mathbf{\Sigma}_2 \tag{65}$$

$$\mathbf{\Delta} = \mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1} + \mathbf{\Sigma}_2 - \mathbf{\Sigma}_Z - \mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1} \tag{66}$$

where clearly $\mathbf{A} \succ \mathbf{0}$, $\mathbf{B} \succeq \mathbf{0}$, and also $\mathbf{\Delta} \succeq \mathbf{0}$ due to Lemma 4. With these selections, we have

$$r(0) = \frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1} + \mathbf{\Sigma}_Z - \mathbf{\Sigma}_2|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1}|} \tag{67}$$

$$r(1) = \frac{1}{2}\log\frac{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1}|}{|\mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1} + \mathbf{\Sigma}_2 - \mathbf{\Sigma}_Z|} \tag{68}$$

Thus, (62) can be expressed as

$$r(1) \leq \alpha \leq r(0) \tag{69}$$

We know from Lemma 7 that $r(t)$ is continuous in t. Then, from the intermediate value theorem, there exists a t^* such that $r(t^*) = \alpha$. Thus, we have

$$\alpha = r(t^*) = \frac{1}{2}\log\frac{|\mathbf{A} + t^*\mathbf{\Delta} + \mathbf{\Sigma}_Z - \mathbf{\Sigma}_2|}{|\mathbf{A} + t^*\mathbf{\Delta}|} \tag{70}$$

$$= \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{71}$$

where $\mathbf{K}^* = \mathbf{A} + t^*\mathbf{\Delta} - \mathbf{\Sigma}_2$. Since $0 \leq t^* \leq 1$, $\mathbf{K}^*$ satisfies the following orderings,

$$\mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1} - \mathbf{\Sigma}_2 \preceq \mathbf{K}^* \preceq \mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1} - \mathbf{\Sigma}_Z \tag{72}$$

which in turn, by using Lemma 2 and Lemma 3, imply the following orderings,

$$\mathbf{K}^* \succeq \mathbf{J}(\mathbf{X}+\mathbf{N}_2)^{-1} - \mathbf{\Sigma}_2 \succeq \mathbf{J}(\mathbf{N}_2)^{-1} - \mathbf{\Sigma}_2 = \mathbf{0} \tag{73}$$

$$\mathbf{K}^* \preceq \mathbf{J}(\mathbf{X}+\mathbf{N}_Z)^{-1} - \mathbf{\Sigma}_Z \preceq \text{Cov}(\mathbf{X}) + \mathbf{\Sigma}_Z - \mathbf{\Sigma}_Z \preceq \mathbf{S} \tag{74}$$

which can be summarized as

$$\mathbf{0} \preceq \mathbf{K}^* \preceq \mathbf{S} \tag{75}$$

In addition, using Lemma 4 in (72), we get $\mathbf{K}^* \succeq \mathbf{J}(\mathbf{X} + \mathbf{N})^{-1} - \mathbf{\Sigma}_N$ for any Gaussian random vector $\mathbf{N}$ such that its covariance matrix satisfies $\mathbf{\Sigma}_N \preceq \mathbf{\Sigma}_2$. This inequality is equivalent to

$$\mathbf{J}(\mathbf{X}+\mathbf{N}) \succeq (\mathbf{K}^* + \mathbf{\Sigma}_N)^{-1}, \qquad \text{for} \qquad \mathbf{\Sigma}_N \preceq \mathbf{\Sigma}_2 \tag{76}$$

where $\mathbf{N}$ is Gaussian with covariance matrix $\mathbf{\Sigma}_N$.

Returning to the proof of Theorem 4, we now lower bound

$$h(\mathbf{X}+\mathbf{N}_Z) - (\mathbf{X}+\mathbf{N}_1) \tag{77}$$

while keeping

$$h(\mathbf{X}+\mathbf{N}_Z) - (\mathbf{X}+\mathbf{N}_2) = \alpha = \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{78}$$

The lower bound on (77) can be obtained as follows

$$h(\mathbf{X}+\mathbf{N}_Z) - h(\mathbf{X}+\mathbf{N}_1) = \frac{1}{2}\int_{\mathbf{\Sigma}_1}^{\mathbf{\Sigma}_Z} \mathbf{J}(\mathbf{X}+\mathbf{N}) d\mathbf{\Sigma}_N \tag{79}$$

$$= \frac{1}{2}\int_{\mathbf{\Sigma}_1}^{\mathbf{\Sigma}_2} \mathbf{J}(\mathbf{X}+\mathbf{N}) d\mathbf{\Sigma}_N + \frac{1}{2}\int_{\mathbf{\Sigma}_2}^{\mathbf{\Sigma}_Z} \mathbf{J}(\mathbf{X}+\mathbf{N}) d\mathbf{\Sigma}_N \tag{80}$$

$$= \frac{1}{2}\int_{\mathbf{\Sigma}_1}^{\mathbf{\Sigma}_2} \mathbf{J}(\mathbf{X}+\mathbf{N}) d\mathbf{\Sigma}_N + \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{81}$$

$$\geq \frac{1}{2}\int_{\mathbf{\Sigma}_1}^{\mathbf{\Sigma}_2} (\mathbf{K}^* + \mathbf{\Sigma}_N)^{-1} d\mathbf{\Sigma}_N + \frac{1}{2}\log\frac{|\mathbf{K}^\star + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{82}$$

$$= \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_2|}{|\mathbf{K}^* + \mathbf{\Sigma}_1|} + \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^* + \mathbf{\Sigma}_2|} \tag{83}$$

$$= \frac{1}{2}\log\frac{|\mathbf{K}^* + \mathbf{\Sigma}_Z|}{|\mathbf{K}^\star + \mathbf{\Sigma}_1|} \tag{84}$$

where (80) follows from the fact that the integral in (79) is path-independent, and (82) is due to Lemma 6 and (76), completing the proof of Theorem 4 for $\mathbf{U} = \phi$.

References

[1] A. Wyner. The wire-tap channel. *Bell System Technical Journal*, 54(8):1355–1387, Jan. 1975.

[2] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Trans. Inf. Theory*, IT-24(3):339–348, May 1978.

[3] A. Khisti, A. Tchamkerten, and G. W. Wornell. Secure broadcasting over fading channels. *IEEE Trans. Inf. Theory*, 54(6):2453–2469, Jun. 2008.

[4] E. Ekrem and S. Ulukus. Secrecy capacity of a class of broadcast channels with an eavesdropper. *EURASIP Journal on Wireless Communications and Networking*. To appear. Also available at [arXiv:0812.0319].

[5] E. Ekrem and S. Ulukus. On secure broadcasting. In *42th Asilomar Conf. Signals, Syst. and Comp.*, Oct. 2008.

[6] G. Bagherikaram, A. S. Motahari, and A. K. Khandani. The secrecy rate region of the broadcast channel. In *46th Annual Allerton Conf. Commun., Contr. and Comput.*, Sep. 2008. Also available at [arXiv:0806.4200].

[7] E. Ekrem and S. Ulukus. Secrecy capacity region of the Gaussian multi-receiver wiretap channel. In *IEEE ISIT*, Jul. 2009.

[8] E. Ekrem and S. Ulukus. The secrecy capacity region of the Gaussian MIMO multi-receiver wiretap channel. Submitted to *IEEE Trans. Inf. Theory*, Mar. 2009. Also available at [arXiv:0903.3096].

[9] A. El Gamal. EE478 Multiple user information theory. Lecture notes.

[10] P. Bergmans. A simple converse for broadcast channels with additive white Gaussian noise. *IEEE Trans. Inf. Theory*, 20(3):279–280, Mar. 1974.

[11] N. M. Blachman. The convolution inequality for entropy powers. *IEEE Trans. Inf. Theory*, IT-11(2):267–271, Apr. 1965.

[12] R. Bustin, R. Liu, H. V. Poor, and S. Shamai. An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel. *EURASIP Journal on Wireless Communications and Networking*. To appear.

[13] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz). The capacity region of the Gaussian multiple-input multiple-output broadcast channel. *IEEE Trans. Inf. Theory*, 52(9):3936–3964, Sep. 2006.

[14] T. Liu and S. Shamai (Shitz). A note on the secrecy capacity of the multi-antenna wiretap channel. *IEEE Trans. Inf. Theory*, 55(6):2547–2553, Jun. 2009.

[15] D. P. Palomar and S. Verdu. Gradient of mutual information in linear vector Gaussian channels. *IEEE Trans. Inf. Theory*, 52(1):141–154, Jan. 2006.

[16] S. Ihara. On the capacity of channels with additive non-Gaussian noise. *Information and Control*, 37(1):34–39, Apr. 1978.

[17] S. H. Diggavi and T. M. Cover. The worst additive noise constraint under a covariance constraint. *IEEE Trans. Inf. Theory*, 47(7):3072–3081, Nov. 2001.

[18] T. Liu and P. Viswanath. An extremal inequality motivated by multiterminal information theoretic problems. *IEEE Trans. Inf. Theory*, 53(5):1839–1851, May 2007.