The Secrecy Capacity of the MIMO Wiretap Channel

Frédérique Oggier and Babak Hassibi

***Abstract*—We consider the MIMO wiretap channel, that is a MIMO broadcast channel where the transmitter sends some confidential information to one user which is a legitimate receiver, while the other user is an eavesdropper. Perfect secrecy is achieved when the transmitter and the legitimate receiver can communicate at some positive rate, while insuring that the eavesdropper gets zero bits of information. In this paper, we compute the perfect secrecy capacity of the multiple antenna MIMO broadcast channel, where the number of antennas is arbitrary for both the transmitter and the two receivers. Our technique involves a careful study of a Sato-like upper bound via the solution of a certain algebraic Riccati equation.**

***Index Terms*—Multiple antennas, secrecy capacity, wiretap channel.**

I. Introduction

SECURITY in wireless communication is a critical issue, which has recently attracted a lot of interest. By nature, wireless channels offer a shared medium, particularly favorable to eavesdropping. Among the numerous points of view from which security has been investigated, we adopt here the one of information theoretic security. In this context, most of the works dealing with wireless communication are based on the seminal work of Wyner [27], and its model, the wire-tap channel.

A. Information Theoretic Confidentiality

In a traditional confidentiality setting, a transmitter (Alice) wants to send some secret message to a legitimate receiver (Bob), and prevent the eavesdropper (Eve) to have knowledge of the message. From an information theoretic point of view, the communication channel involved can be modeled as a broadcast channel, following the wire-tap channel model introduced by Wyner [27]: a transmitter broadcasts its message, say w^k encoded into a codeword x^n, and the two receivers (the legitimate and the illegitimate) respectively receive y^n and z^n, the output of their channel. The knowledge that the eavesdropper gets of w^k from its received signal z^n is modeled by

$$I(z^n; w^k) = h(w^k) - h(w^k \,|\, z^n) \tag{1}$$

since the mutual information measures the amount of information that z^n contains about w^k. The notion of perfect secrecy captures the idea that whatever are the resources available to the eavesdropper, they will not allow him to get a single bit of information. Perfect secrecy thus requires

$$I(z^n; w^k) = 0 \Longleftrightarrow h(w^k) = h(w^k \,|\, z^n) \tag{2}$$

to hold asymptotically as n grows. In other words, the amount of randomness is the same in w^k or in $w^k \,|\, z^n$. The decoder computes an estimate $\hat{w}^k$ of the transmitted message w^k, and the probability P_e of decoding erroneously is given by

$$P_e = Pr(w^k \neq \hat{w}^k). \tag{3}$$

The amount of ignorance that the eavesdropper has about a message w^k is called the *equivocation rate*, and following the above discussion, it is naturally defined as:

Definition 1: The *equivocation rate* R_e at the eavesdropper is

$$R_e = \frac{1}{n} h(w^k \,|\, z^n) \tag{4}$$

with $0 \leq R_e \leq h(w^k)/n$.
Clearly, if R_e is equal to the information rate $h(w^k)/n$, then $I(z^n; w^k) = 0$, which yields perfect secrecy. Associated with secrecy is a perfect secrecy rate R_S, which is the amount of information that can be sent not only reliably but also confidentially, with the help of a $(2^{nR_S}, n)$ code.

Definition 2: A *perfect secrecy rate* R_S is said to be achievable if for any $\epsilon, \epsilon' > 0$, there exists a sequence of $(2^{nR_S}, n)$ codes such that for any $n \geq n(\epsilon, \epsilon')$, we have

$$P_e \leq \epsilon' \tag{5}$$

$$R_S - \epsilon \leq R_e. \tag{6}$$

The first condition (5) is the standard definition of achievable rate as far as reliability is concerned. The second condition (6) guarantees secrecy, up to the equivocation rate, which we will require to be $h(w^k)/n$ to have perfect secrecy. The secrecy capacity is defined similarly to the standard capacity:

Definition 3: The *secrecy capacity* C_S is the maximum achievable perfect secrecy rate.

Manuscript received July 19, 2008; revised September 25, 2010; accepted March 18, 2011. Date of current version July 29, 2011. This work was done while F. Oggier was with the Department of Electrical Engineering, California Institute of Technology, Pasadena. This work was supported in part by NSF Grant CCF-0729203, in part by Caltech's Lee Center for Advanced Networking, and in part by a grant from the David and Lucille Packard Foundation. Part of this work appeared at the 45th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, July 2007, and the IEEE International Symposium on Information Theory, Toronto, ON, Canada, July 2008.

F. Oggier is with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore (e-mail: frederique@ntu.edu.sg).

B. Hassibi is with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: hassibi@systems.caltech.edu).

Communicated by S. Ulukus, Associate Editor for Communication Networks.

Digital Object Identifier 10.1109/TIT.2011.2158487

B. Previous Work

In his seminal work [27], Wyner showed for discrete memoryless channels that the perfect secrecy capacity is actually the difference of the respective mutual informations of the two users, maximized over the input distribution. To prove this result, he worked under the assumption that the channel of the eavesdropper is a degraded version of the channel of the legitimate receiver. This result has been generalized to Gaussian channels by Leung *et al.* [12], under the same assumption.

The wire-tap channel has been adopted as a model for numerous works on information theoretic security, and in particular for those on fading channels, both for point-to-point and multi-user systems. We mainly review the prior work for the point-to-point case. In [6], Gopala *et al.* have shown that the secrecy capacity is also the difference of the two mutual informations maximized over the input distribution in the case of a single antenna fading channel, under the assumption of asymptotically long coherence intervals, when the transmitter either knows both channels or only the legitimate channel. When only the legitimate channel is known, an optimal power allocation is given, using a variable rate transmission scheme. In [1], Barros *et al.* have characterized information theoretic security in terms of outage probability. In the case when the transmitter does not know the eavesdropper channel, they define and compute the probability of transmitting at a secrecy rate R_S bigger than the secrecy capacity C_S (i.e., the outage probability) as the probability that the information theoretic security is compromised. They also show that the probability that the secrecy capacity C_S is positive can actually be greater than zero even if the average SNR of the legitimate channel is weaker than the one of the eavesdropper. They extend their work in [2], where they also consider the cases when Alice has either imperfect or perfect knowledge of the eavesdropper channel. Independently, Liang *et al.* [15] and Li *et al.* [13] have computed the secrecy capacity for the parallel wiretap channel with independent subchannels, and derived optimal source power allocation. The secrecy capacity of the wiretap channel with single antenna fading channel follows. Finally, the results of [15] are extended in [16], where a fading broadcast channel with confidential messages is considered, with common information for two receivers, and confidential information intended for only one receiver. The secrecy capacity is computed for the parallel broadcast channel with both independent and degraded subchannels.

In this work, we are interested in the perfect secrecy capacity of multiple antenna channels. A first study of the problem has been proposed by Hero [7]. In a different context than the wire-tap channel, he introduced the so-called constraints of low probability of detection, and low probability of intercept, considering the scenario where the transmitter and the receiver are both informed about their channel while the eavesdropper is uniformed about his. In [23], the SIMO wiretap channel has been considered. Several results on the secrecy in MIMO communication have been provided recently. In [14], lower bounds on the secrecy capacity is computed for the MISO case. Furthermore, a lower bound is computed in the MIMO case. This lower bound, which proves an achievability result for the secrecy capacity, is shown to be the expected result, namely, the difference of the two user mutual informations maximized over the input distribution, like in the previous cases. The secrecy capacity for the MISO case has also been proven independently by Khisti *et al.* [8] and Shafiee *et al.* [24]. In [8], the authors furthermore give an upper bound for the MIMO case, in a regime asymptotic in SNR. The secrecy capacity has been computing for the particular cases where both the transmitter and receiver have two antennas, and the eavesdropper has either one antenna [25] or two antennas [21]. Finally, Liu *et al.* [17], [18] computed the secrecy capacity for a Gaussian broadcast channel, where a multi-antenna transmitter sends independent confidential messages to two users.

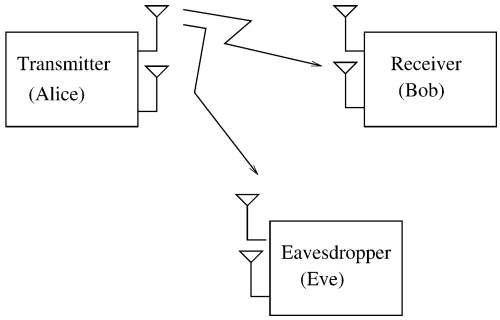

Fig. 1. MIMO wiretap channel.

The contribution of this paper is to compute the perfect secrecy capacity of the multiple antenna wire-tap channel, for any number of transmit/receive antennas, as well as for any SNR regime. One of the difficulties in studying the MIMO wire-tap channel is that the broadcast MIMO channel is not degraded, an assumption which is crucial in the proof of the converse in the original paper by Wyner (as well as in the proofs presented in [12], [6], [1], [15]). In order to compute the secrecy capacity, we provide a proof technique for the converse, which is different from the original one, and allows us to deal with channels that are not degraded. Note that our result shows that the inner bound by Li *et al.* [14] is tight, and this is proved by the computation of an upper bound that actually matches the lower bound. Independently of our results, Khisthi and Wornell [9]–[11] have also computed the secrecy capacity of the MIMO wiretap channel (which they refer to as MIMO-ME). They use alternative means, namely they exploit Czisár and Körner's result [4] and identify the necessary auxiliary random variables. Furthermore, their characterization is through a saddlepoint (essentially, the optimization in Proposition 5 below), whereas our characterization is through a single optimization (see Theorem 1 below). An alternative derivation of our result, and that of Khisti-Wornell, has also appeared in Liu and Shamai [19]. Finally, a characterization of the secrecy capacity using a MMSE approach has appeared in [3].

C. MIMO Wiretap Channel

We consider the MIMO wiretap channel (see Fig. 1), that is, a broadcast channel where the transmitter is equipped with n transmit antennas, while the legitimate receiver and an eavesdropper have respectively n_M and n_E receive antennas. Thus, our model is described by the following broadcast channel:

$$Y = H_M X + V_M \tag{7}$$

$$Z = H_E X + V_E \tag{8}$$

where Y, V_M, and Z, V_E are respectively $n_M \times 1$ and $n_E \times 1$ vectors. The notation that we will use throughout the paper is that the subscript M refers to the main channel (the one of the legitimate receiver), while the subscript E refers to the eavesdropper channel. We will denote by $\mathbf{I}_n$ the $n \times n$ identity matrix, and by $\mathbf{0}_n$ the $n \times n$ all zero matrix. We may omit the subscript if the dimension is obvious. We make the following assumptions:

- X is the $n \times 1$ complex transmitted signal, with $K_X = E[XX^*] \succeq \mathbf{0}_n$ satisfying the power constraint

$$Tr(K_X) \leq P. \tag{9}$$

  This power constraint holds for the whole paper, and we may sometimes omit it. Note though that the optimizations that follow will instead use as power constraint $Tr(K_X) = P$ instead of (9), since we can fix the power to $P' \leq P$, solve the optimization for P', and then pick the optimal P' which is P.
- H_M and H_E are respectively $n_M \times n$ and $n_E \times n$ fixed channel matrices. They are both assumed to be known at the transmitter and receiver. Along the paper we will usually consider two cases: the *definite case*, that is when $H_M^* H_M \succ H_E^* H_E$ or $H_E^* H_E \succ H_M^* H_M$, which corresponds to the degraded case, and the *indefinite case*, which is when some of the eigenvalues of $H_E^* H_E - H_M^* H_M$ are positive, and other negative or zero. We discard the meaningless case where $H_M^* H_M = H_E^* H_E$.
- V_M, V_E are independent circularly symmetric complex Gaussian vectors with identity covariance $K_M = \mathbf{I}_{n_M}$, $K_E = \mathbf{I}_{n_E}$ and independent of the transmitted signal X.

Theorem 1: The secrecy capacity of the MIMO wiretap channel is given by

$$C_S = \max_{K_X \succeq \mathbf{0}, Tr(K_X)=P} \log\det(\mathbf{I} + H_M K_X H_M^*) - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{10}$$

The paper contains the proof of the above theorem: in Section II, we prove an achievability result which characterizes the optimal matrices $\tilde{K}_X$, while Section 3 contains the main results, namely the proof of the converse.

Since our proof uses mainly optimization techniques, we may use some well known optimization facts without explicit references. The reader may refer to [20] when no explicit other reference is given.

II. On the Achievability

In this section, we state the achievability part of the secrecy capacity, and further prove that in the nondegraded case, the achievability is maximized by $n \times n$ matrices K_X which are low rank, that is of any rank $r < n$.

Proposition 1: The perfect secrecy rate

$$R_S = \max_{K_X \succeq \mathbf{0}, Tr(K_X)=P} \log\det(\mathbf{I} + H_M K_X H_M^*) - \log\det(\mathbf{I} + H_E K_X H_E^*) \tag{11}$$

is achievable.

This has already been proved [14].[1] In fact, the interpretation is obvious. When K_X is chosen, the difference between the resulting mutual informations to the legitimate user and eavesdropper can be secretly transmitted.

Proposition 2: Let $\tilde{K}_X$ be an optimal solution to the optimization problem

$$\max_{K_X \succeq \mathbf{0}, Tr(K_X)=P} \log\det(\mathbf{I} + H_M K_X H_M^*) - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{12}$$

1) If $H_E^* H_E - H_M^* H_M$ is either indefinite or semidefinite, then $\tilde{K}_X$ is a low rank matrix.
2) Let r be the rank of $\tilde{K}_X$. If $n > n_M$, then for at least one optimal solution we have that $r \leq n_M$.

Proof: 1) In order to show that the optimal $\tilde{K}_X$ is low rank, we define a Lagrangian which includes the power constraint, but not the non-negativity constraint $K_X \succeq \mathbf{0}$, and show that this yields no solution. From there, we can conclude that the optimal solution is on the boundary of the cone of positive semi-definite matrices, namely matrices of rank $r < n$. We thus define the following Lagrangian:

$$\log\det(\mathbf{I}_{n_M} + H_M K_X H_M^*) - \log\det(\mathbf{I}_{n_E} + H_E K_X H_E^*) - \lambda Tr(K_X) \tag{13}$$

and look for its stationary points, that is for the solution of the following equation:

$$\nabla_{K_X}(\log\det(\mathbf{I} + H_M K_X H_M^*) - \log\det(\mathbf{I} + H_E K_X H_E^*) - \lambda Tr(K_X)) = 0 \iff H_M^* H_M (\mathbf{I} + K_X H_M^* H_M)^{-1} = (\mathbf{I} + H_E^* H_E K_X)^{-1} H_E^* H_E + \lambda \mathbf{I}_n. \tag{14}$$

By premultiplying the above equation by $(\mathbf{I} + H_E^* H_E K_X)$ and postmultiplying it by $(\mathbf{I} + K_X H_M^* H_M)$, we get

$$H_M^* H_M - H_E^* H_E = \lambda(\mathbf{I} + H_E^* H_E K_X)(\mathbf{I} + K_X H_M^* H_M). \tag{15}$$

If $\lambda = 0$, then $H_M^* H_M = H_E^* H_E$ which is discarded by our model assumption. Thus, equivalently, by further pre and postmultiplying by K_X, we can rewrite

$$K_X(H_M^* H_M - H_E^* H_E)K_X \tfrac{1}{\lambda} = (K_X + K_X H_E^* H_E K_X)(K_X + K_X H_M^* H_M K_X). \tag{16}$$

Now if we assume that $K_X \succ \mathbf{0}$, then all the eigenvalues of

$$(K_X + K_X H_E^* H_E K_X)(K_X + K_X H_M^* H_M K_X) \tag{17}$$

are strictly positive (see Lemma 1 below). This implies that the above equation can have a solution if and only if the Hermitian matrix

$$K_X(H_M^* H_M - H_E^* H_E)K_X \frac{1}{\lambda} \tag{18}$$

[1] As pointed out by a reviewer, this proposition also follows from either Wyner's achievability result or Csiszár–Körner's capacity formula by an appropriate selection of the underlying random variables.

is positive definite. This means that either $H_M^*H_M \succ H_E^*H_E$ and $\lambda > 0$, or $H_M^*H_M \prec H_E^*H_E$ and $\lambda < 0$. This gives a contradiction if $H_M^*H_M - H_E^*H_E$ is either indefinite or semidefinite, implying that K_X has to be low rank.

2) Let us now assume that $n > n_M$ (we make no assumption on whether the channel is degraded). Let r be the rank of K_X, so that we can decompose K_X as $K_X = U_X U_X^*$ for some $n \times r$ matrix U_X. We now define a Lagrangian similarly as above, namely

$$\begin{array}{c} \log\det(\mathbf{I}_n + U_X U_X^* H_M^* H_M) \\ - \log\det(\mathbf{I}_n + U_X U_X^* H_E^* H_E) \\ -\lambda(Tr(U_X U_X^*) - P). \end{array} \tag{19}$$

Note that $\lambda \geq 0$, since the maximization over U_X of the considered objective function subject to the constraint that $Tr(U_X U_X^*) \leq P$ is equivalent to

$$\begin{array}{c} \max_{U_X} \min_{\lambda \geq 0} \log\det(\mathbf{I}_n + U_X U_X^* H_M^* H_M) \\ - \log\det(\mathbf{I}_n + U_X U_X^* H_E^* H_E) - \lambda(Tr(U_X U_X^*) - P). \end{array} \tag{20}$$

We now compute the derivative of the Lagrangian, which yields

$$\begin{aligned} & 2U_X^* H_M^* H_M (\mathbf{I}_n + U_X U_X^* H_M^* H_M)^{-1} \\ &= \lambda U_X^* + 2U_X^* H_E^* H_E (\mathbf{I}_n + U_X U_X^* H_E^* H_E)^{-1} && (21) \\ &= U_X^* (\lambda \mathbf{I}_n + 2H_E^* H_E (\mathbf{I}_n + U_X U_X^* H_E^* H_E)^{-1}) && (22) \\ &= U_X^* (\lambda \mathbf{I}_n + 2H_E^* (\mathbf{I}_{n_E} + H_E U_X U_X^* H_E^*)^{-1} H_E). && (23) \end{aligned}$$

The rank of the matrix on the left is smaller or equal to $\min(r, n, n_M)$.

When $\lambda > 0$, the matrix on the right hand side has rank r (because the second matrix is strictly positive definite: it is the sum of the identity matrix and a matrix which is itself strictly positive definite) which means that

$$r \leq \min(r, n, n_M) = \min(r, n_M) \tag{24}$$

which yields $r \leq n_M$ as desired.

When $\lambda = 0$, the derivative of the Lagrangian simplifies to

$$\begin{array}{c} U_X^* H_M^* H_M (\mathbf{I}_n + U_X U_X^* H_M^* H_M)^{-1} \\ = U_X^* H_E^* H_E (\mathbf{I}_n + U_X U_X^* H_E^* H_E)^{-1} \end{array} \tag{25}$$

or equivalently

$$\begin{array}{c} (\mathbf{I}_n + U_X^* H_M^* H_M U_X)^{-1} U_X^* H_M^* H_M (\mathbf{I}_n + U_X U_X^* H_E^* H_E) \\ = U_X^* H_E^* H_E \end{array} \tag{26}$$

which yields

$$U_X^* (H_E^* H_E - H_M^* H_M) = \mathbf{0}. \tag{27}$$

This clearly implies that the objective function is zero, since the quantities inside the two logs become identical. But $U_X = \mathbf{0}$ also makes the objective function zero, so we might as well take $U_X = \mathbf{0}$, for which $r = 0 \leq n_M$ as desired. ∎

Note that this result is intuitively clear, since the transmitter should not use the directions in which the eavesdropper is stronger, which explains why the optimal matrix K_X is low rank.

Lemma 1: If $A = A^* \succ \mathbf{0}$ and $B = B^* \succ \mathbf{0}$, then the matrix AB has all positive eigenvalues.

Proof: Since $A \succ \mathbf{0}$, we can write $A = A^{1/2}(A^*)^{1/2}$ with $A^{1/2}$ invertible. Therefore

$$AB = A^{1/2}((A^*)^{1/2} B A^{1/2}) A^{-1/2} \tag{28}$$

has the same eigenvalues as the matrix $(A^*)^{1/2} B A^{1/2}$, which is positive definite. ∎

III. Proof of the Converse

The goal of this section is to prove the converse, namely

Theorem 2: For any sequence of $(2^{nR_S}, n)$ codes with probability of error $P_e \leq \epsilon'$ and equivocation rate $R_S - \epsilon \leq R_e$ for any $n \geq n(\epsilon, \epsilon')$, $\epsilon, \epsilon' > 0$, then the secrecy rate R_S satisfies

$$\begin{array}{c} R_S \leq \max_{K_X \succeq \mathbf{0}, Tr(K_X) = P} \log\det(\mathbf{I} + H_M K_X H_M^*) - \\ \log\det(\mathbf{I} + H_E K_X H_E^*). \end{array} \tag{29}$$

The proof is done in three main steps, that we briefly sketch before entering into the details.

Step 1. We have, similarly to [12], [6] that

$$R_S - \epsilon \leq \frac{1}{n}[I(X^n; Y^n \mid Z^n) + \delta], \epsilon, \delta > 0. \tag{30}$$

Thus, all the work consists of maximizing $I(X; Y \mid Z)$. In Subsection 3.1, we start by proving the following upper bound:

$$I(X; Y \mid Z) \leq \max_{K_X \succeq \mathbf{0}} \tilde{I}(X; Y \mid Z) \tag{31}$$

where

$$\begin{array}{c} \tilde{I}(X; Y | Z) = \\ \log\det\left[\mathbf{I}_n + (H_M^* \; H_E^*) \begin{bmatrix} \mathbf{I}_{n_M} & A \\ A^* & \mathbf{I}_{n_E} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X\right] \\ - \log\det(\mathbf{I} + H_E K_X H_E^*) \end{array} \tag{32}$$

and A is an $n_M \times n_E$ matrix which denotes the correlation between V_M and V_E. Note that clearly

$$\begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{33}$$

and strict inequality is required for the inverse to exist. At this point of the proof, the converse can be proved for the two "simple" cases when $H_M^*H_M \succ H_E^*H_E$ and $H_E^*H_E \succ H_M^*H_M$, which are the cases when the channel is degraded.

Step 2. In practice, V_M and V_E are independent. However, the secrecy capacity does not depend on A: the legitimate receiver and eavesdropper have no way of exploiting any possible correlation, since they decode based on their separate received signals. This is an argument classically used for broadcast channels. Thus, we can assume that $\tilde{I}(X; Y \mid Z)$ is a function of both A and K_X for the purposes of tightening our upper bound. Since

$$I(X; Y \mid Z) \leq \max_{K_X \succeq \mathbf{0}} \tilde{I}(X; Y \mid Z) \tag{34}$$

for all A such that $\mathbf{I} - AA^* \succ \mathbf{0}_{n_E}$, we have that

$$I(X; Y \mid Z) \leq \min_{A} \max_{K_X \succeq \mathbf{0}} \tilde{I}(X; Y \mid Z). \tag{35}$$

By proving (in Subsection 3.2) that $\tilde{I}(X;Y|Z)$ is actually concave in K_X and convex in A, we are then allowed to write that

$$I(X;Y\,|\,Z) \leq \max_{K_X \succeq \mathbf{0}} \min_{A} \tilde{I}(X;Y\,|\,Z). \tag{36}$$

We can now jointly optimize $\tilde{I}(X;Y\,|\,Z)$ over K_X and A, and compute the optimal $\tilde{A}$ in closed form expression, while showing that the optimal $\tilde{K}_X$ is on the boundary of its domain, namely, $\tilde{K}_X$ is low rank.

Step 3. We conclude the proof (Subsection 3.3) by showing that the converse matches the achievability, using the closed form expression for the optimal $\tilde{A}$.

A. Bound on $I(X;Y\,|\,Z)$ and Result for the Degraded Case

We start by recalling a standard result, which has already been proved in [12], [6].

Lemma 2: Given any sequence of $(2^{nR_S}, n)$ codes with $P_e \leq \epsilon'$ and $R_S - \epsilon \leq R_e$ for any $n \geq n(\epsilon,\epsilon')$, $\epsilon,\epsilon' > 0$, the secrecy rate R_S can be upper bounded as follows:

$$R_S - \epsilon \leq \frac{1}{n}[I(X;Y\,|\,Z) + \delta] \tag{37}$$

for $\epsilon, \delta > 0$.

We thus focus now on finding an upper bound on $I(X;Y\,|\,Z)$. We provide two approaches:

1) An upper bound is given by assuming that the legitimate receiver knows both his channel and the one of the eavesdropper.
2) The same upper bound can also be obtained as follows. Clearly, $I(X;Y\,|\,Z)$ is upper bounded by taking the supremum over all input distributions $\mathcal{P}(X)$

$$I(X;Y\,|\,Z) \leq \sup_{\mathcal{P}(X)} I(X;Y\,|\,Z) = \max_{K_X \succeq \mathbf{0}} \tilde{I}(X;Y\,|\,Z) \tag{38}$$

where $\tilde{I}(X;Y\,|\,Z)$ denotes the value of $I(X;Y\,|\,Z)$ when $\mathcal{P}(X)$ is optimal. We will prove that the optimal distribution is Gaussian.

Proposition 3: We have the following upper bound:

$$\begin{aligned} I(X;Y\,|\,Z) \leq \max_{K_X \succeq \mathbf{0}} \log\det\Big[\mathbf{I}_n + \\ (H_M^* \; H_E^*) \begin{bmatrix} \mathbf{I}_{n_M} & A \\ A^* & \mathbf{I}_{n_E} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X \Big] \\ - \log\det(\mathbf{I} + H_E K_X H_E^*) \end{aligned} \tag{39}$$

where A denotes the correlation between V_M and V_E and satisfies $\mathbf{I} - AA^* \succ \mathbf{0}$.

Proof: **First approach**: An upper bound on $I(X;Y\,|\,Z)$ is obtained by assuming that the legitimate receiver knows both its channel H_M and the one of the eavesdropper H_E. In this case, the capacity of the link between the transmitter and the legitimate receiver is that of a MIMO system, namely

$$\begin{aligned} \max_{K_X \succeq \mathbf{0}} \log\det\Big[\mathbf{I}_n + \\ (H_M^* \; H_E^*) \begin{bmatrix} \mathbf{I}_{n_M} & A \\ A^* & \mathbf{I}_{n_E} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X \Big]. \end{aligned} \tag{40}$$

Now the channel we consider is degraded, and an upper bound is thus the difference of the respective mutual informations of the two users maximized over the input distribution, which yields the result.

Second Approach: We now provide an alternative proof. Clearly

$$I(X;Y\,|\,Z) \leq \sup_{\mathcal{P}(X)} I(X;Y\,|\,Z) \tag{41}$$

where $\mathcal{P}(X)$ denotes the input distribution. Now note that

$$\begin{aligned} I(X;Y\,|\,Z) &= h(Y|Z) - h(Y|X,Z) && (42)\\ &= h(Y\,|\,Z) - h(X,Y,Z) + h(X,Z) && (43)\\ &= h(Y\,|\,Z) - h(X) \\ &\quad - h(Y,Z\,|\,X) + h(X) + h(Z\,|\,X) && (44)\\ &= h(Y\,|\,Z) - h(V_E,V_M) + h(V_E). && (45) \end{aligned}$$

Since we know that under second order statistics $h(Y\,|\,Z)$ is maximized by letting (Y,Z) be jointly Gaussian,[2] the optimal is given by choosing X Gaussian. Thus, we have that

$$\begin{aligned} &I(X;Y\,|\,Z) \\ &= h(Y|Z) - h(V_E,V_M) + h(V_E) && (46)\\ &= h(Y,Z) - h(Z) - h(V_E,V_M) + h(V_E) && (47) \end{aligned}$$

which, when X is Gaussian, is given by

$$\begin{aligned} &\log\det(K_{YZ}) - \log\det(K_Z) \\ &- \log\det(K_{ME}) + \log\det(K_E) \end{aligned} \tag{48}$$

where K_{YZ}, K_Z, K_{ME} and $K_E = \mathbf{I}_{n_E}$ are covariance matrices, with

$$K_{YZ} = \begin{bmatrix} H_M K_X H_M^* + \mathbf{I}_{n_M} & H_M K_X H_E^* + A \\ H_E K_X H_M^* + A^* & H_E K_X H_E^* + \mathbf{I}_{n_E} \end{bmatrix} \tag{49}$$

where A denotes the correlation between V_M and V_E, and

$$K_{ME} = \begin{bmatrix} \mathbf{I}_{n_M} & A \\ A^* & \mathbf{I}_{n_E} \end{bmatrix}. \tag{50}$$

In order for K_{ME} to be well defined, A has to satisfy $\mathbf{I} - AA^* \succeq \mathbf{0}$. Thus, we have

$$\begin{aligned} &\log\det\left[\begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix} + \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X (H_M^* \; H_E^*)\right] \\ &\quad - \log\det(H_E K_X H_E^* + \mathbf{I}) - \log\det(K_{ME}) \\ &= \log\det\left[\mathbf{I} + \begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X (H_M^* \; H_E^*)\right] \\ &\quad - \log\det(H_E K_X H_E^* + \mathbf{I}) \end{aligned} \tag{51}$$

where the second equality is well defined if we further require $\mathbf{I} - AA^* \succ \mathbf{0}$. The value of $I(X;Y\,|\,Z)$ when X is Gaussian is denoted by $\tilde{I}(X;Y\,|\,Z)$

$$\begin{aligned} &\tilde{I}(X;Y|Z) = \\ &\log\det\left[\mathbf{I} + (H_M \; H_E^*) \begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} K_X\right] \\ &\quad - \log\det(\mathbf{I} + H_E K_X H_E^*). \end{aligned} \tag{52}$$

■

[2]This is known as the Thomas inequality [26]. The authors would like to thank Prof. Shlomo Shamai for pointing out this argument.

We can now conclude the proof of the converse for the "simple" cases when $H_M^* H_M \succ H_E^* H_E$ or $H_E^* H_E \succ H_M^* H_M$.[3]

Proposition 4:

1) If $H_M^* H_M \succ H_E^* H_E$, we have that

$$I(X;Y|Z) \leq \max_{K_X \succ \mathbf{0}} \log\det(\mathbf{I} + H_M K_X H_M^*) - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{53}$$

2) Vice versa, if $H_E^* H_E \succ H_M^* H_M$, we have that

$$I(X;Y \mid Z) = 0. \tag{54}$$

Proof: Let us first introduce two other ways of writing $\tilde{I}(X;Y \mid Z)$, as defined in (52). Note first the following UDL factorization:

$$\begin{bmatrix} \mathbf{I}_{n_M} & A \\ A^* & \mathbf{I}_{n_E} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & A \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} - AA^* & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ A^* & \mathbf{I} \end{bmatrix} \tag{55}$$

so that

$$\begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -A^* & \mathbf{I} \end{bmatrix} \begin{bmatrix} (\mathbf{I} - AA^*)^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & A \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \tag{56}$$

and we have that

$$(H_M^* \; H_E^*) \begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} = (H_M^* - H_E^* A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E) + H_E^* H_E. \tag{57}$$

A 1st equivalent formula for $I(X;Y \mid Z)$ is now given by

$$\tilde{I}(X;Y \mid Z) = \log\det(\mathbf{I} + ((H_M^* - H_E^* A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E) + H_E^* H_E)K_X) - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{58}$$

By considering now a LDU factorization, we get

$$\begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ A^* & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} - AA^* \end{bmatrix} \begin{bmatrix} \mathbf{I} & A \\ \mathbf{0} & \mathbf{I} \end{bmatrix}. \tag{59}$$

Thus

$$\begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{I} & -A \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I} - A^* A)^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -A^* & \mathbf{I} \end{bmatrix} \tag{60}$$

so that

$$(H_M^* \; H_E^*) \begin{bmatrix} \mathbf{I} & A \\ A^* & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} H_M \\ H_E \end{bmatrix} = (H_E^* - H_M^* A)(\mathbf{I} - A^* A)^{-1}(H_E - A^* H_M) + H_M^* H_M \tag{61}$$

and a 2nd equivalent formula for $I(X;Y \mid Z)$ is given by

$$\tilde{I}(X;Y|Z) = \log\det(\mathbf{I} + ((H_E^* - H_M^* A)(\mathbf{I} - A^* A)^{-1}(H_E - A^* H_M) + H_M^* H_M)K_X) - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{62}$$

[3]As pointed out by one reviewer, an alternative proof is to use Wyner's result and then the worst noise result of Diggavi and Cover to show that a Gaussian input is optimal.

Since the secrecy capacity does not depend on the noise correlation A, and that

$$I(X;Y \mid Z) \leq \max_{K_X} \tilde{I}(X;Y \mid Z) \tag{63}$$

for all A such that $\mathbf{I} - AA^* \succ \mathbf{0}$, we are free to take any such A which does not depend on a choice of K_X.

1) We consider first the case where $H_M^* H_M \succ H_E^* H_E$. By hypothesis $H_M^* H_M$ is invertible (clearly $H_E^* H_E \succeq \mathbf{0}$) and we choose $A^* = H_E(H_M^* H_M)^{-1} H_M^*$. Such A does not depend on a choice of K_X and satisfies $\mathbf{I} - A^* A \succ \mathbf{0}$ (or equivalently $\mathbf{I} - AA^* \succ \mathbf{0}$) since

$$\mathbf{I} - A^* A = \mathbf{I} - H_E(H_M^* H_M)^{-1} H_E^* \succ \mathbf{0} \tag{64}$$

by the degradedness assumption. By plugging A^* into (62), we see that

$$-A^* H_M + H_E = \mathbf{0} \tag{65}$$

which yields the desired result.

2) Similarly if $H_E^* H_E \succ H_M^* H_M$, we choose $A^* = H_E(H_E^* H_E)^{-1} H_M^*$, which satisfies

$$\mathbf{I} - AA^* = \mathbf{I} - H_M(H_E^* H_E)^{-1} H_M^* \succ \mathbf{0}. \tag{66}$$

Since $H_M^* - H_E A^* = \mathbf{0}$, we see from (58) that

$$\tilde{I}(X;Y \mid Z) = 0. \tag{67}$$

■

The cases described in the above proposition can be understood as a simple generalization of the scalar case, since those are the degraded cases. When $H_M^* H_M \succ H_E^* H_E$, the legitimate receiver has a stronger channel in every possible spatial direction, and the capacity is given by the difference of the respective mutual informations of the two users maximized over the input distribution, while if $H_E^* H_E \succ H_M^* H_M$, then this is true for the eavesdropper, and thus, no positive secrecy capacity can be achieved.

We are now left with the case when $H_M^* H_M - H_E^* H_E$ is indefinite, which is the nondegraded case, and thus the interesting case to understand.

B. Minimization Over A and Maximization Over K_X

We have shown in Proposition 3 that

$$I(X;Y \mid Z) \leq \max_{K_X \succeq \mathbf{0}} \tilde{I}(X;Y \mid Z). \tag{68}$$

Since this is true for all A such that $\mathbf{I} - AA^* \succ \mathbf{0}$, we further have that

$$I(X;Y \mid Z) \leq \min_{A} \max_{K_X \succeq \mathbf{0}} \tilde{I}(X;Y \mid Z). \tag{69}$$

To understand this double optimization, we start by analyzing the function $\tilde{I}(X;Y \mid Z)$.

Proposition 5: The function $\tilde{I}(X;Y \mid Z)$ defined in (52) is concave in K_X and convex in A. Consequently

$$\min_{A}\max_{K_X}\tilde{I}(X;Y \mid Z) = \max_{K_X}\min_{A}\tilde{I}(X;Y \mid Z) \tag{70}$$

where K_X and A respectively satisfy $Tr(K_X) = P$, $K_X \succeq \mathbf{0}$, $\mathbf{I} - AA^* \succ \mathbf{0}$.

Proof: Recall from (52) that $\tilde{I}(X;Y \mid Z)$ is given by

$$\log\det\left[\mathbf{I} + (H_M\ H_E^*)\begin{bmatrix}\mathbf{I} & A\\ A^* & \mathbf{I}\end{bmatrix}^{-1}\begin{bmatrix}H_M\\ H_E\end{bmatrix}K_X\right] - \log\det(\mathbf{I} + H_E K_X H_E^*). \tag{71}$$

1) **Convexity in A.** Set

$$C := \begin{bmatrix}\mathbf{I} & A\\ A^* & \mathbf{I}\end{bmatrix}, D := \begin{bmatrix}H_M\\ H_E\end{bmatrix}K_X(H_M^*, H_E^*). \tag{72}$$

Now $\tilde{I}(X;Y \mid Z)$ is of the form $\log\det(I_{n_M+n_E} + C^{-1}D)$, plus some constant term, where $D \succeq \mathbf{0}$. It is well known that $\log\det(\mathbf{I}+C^{-1}D)$ is convex in C. Furthermore, it is convex in any block of C, hence, convex in A. Finally, the set of A such that $\mathbf{I} - AA^* \succ \mathbf{0}$ is convex.

2) **Concavity in K_X.** Recall from (58) that

$$\begin{aligned}&\tilde{I}(X;Y|Z) =\\ &\log\det(\mathbf{I} + ((H_M^* - H_E^*A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E)\\ &\qquad + H_E^*H_E)K_X)\\ &\qquad - \log\det(\mathbf{I} + H_E K_X H_E^*).\end{aligned} \tag{73}$$

Set

$$B := (H_M^* - H_E^*A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E) + H_E^*H_E. \tag{74}$$

We thus have that $\tilde{I}(X;Y \mid Z)$ is given by

$$\log\det(\mathbf{I}_n + BK_X) - \log\det(\mathbf{I}_n + H_E^*H_EK_X) \tag{75}$$

with $B \succeq H_E^*H_E$. We now compute its gradient with respect to K_X

$$B(\mathbf{I} + K_XB)^{-1} - H_E^*H_E(\mathbf{I} + K_XH_E^*H_E)^{-1}$$

which is also equal to

$$B(\mathbf{I} + K_XB)^{-1} - (\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E \succeq \mathbf{0} \tag{76}$$

since $B \succeq H_E^*H_E$ implies that

$$(\mathbf{I} + H_E^*H_EK_X)B \succeq H_E^*H_E(\mathbf{I} + K_XB).$$

Recall that

$$\frac{\partial(X^{-1})_{kl}}{\partial X_{ij}} = -(X^{-1})_{ki}(X^{-1})_{jl} \tag{77}$$

so that the derivative of $F := (\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E$ is a $n^2 \times n^2$ matrix given by

$$\begin{aligned}&\begin{bmatrix}-FF_{11} & -FF_{12} & \dots & -FF_{1n}\\ -FF_{21} & -FF_{22} & \dots & -FF_{2n}\\ \vdots & & & \vdots\\ -FF_{n1} & -FF_{n2} & & -FF_{nn}\end{bmatrix}\\ &= -(\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E\otimes\\ &\quad(\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E.\end{aligned} \tag{78}$$

To check the concavity in K_X, we are thus left to check that

$$\begin{aligned}&(\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E\otimes\\ &(\mathbf{I} + H_E^*H_EK_X)^{-1}H_E^*H_E\\ &\preceq B(\mathbf{I} + K_XB)^{-1}\otimes B(\mathbf{I} + K_XB)^{-1}\end{aligned} \tag{79}$$

which is true by (76).

3) Since we have shown above that $\tilde{I}(X;Y \mid Z)$ is concave in K_X and convex in A, we have a saddle point and, therefore, [20]

$$\min_{A}\max_{K_X}\tilde{I}(X;Y \mid Z) = \max_{K_X}\min_{A}\tilde{I}(X;Y \mid Z). \tag{80}$$

■

By combining the above result with (69), we can exchange the order of the two optimizations, to get

$$I(X;Y \mid Z) \le \max_{K_X}\min_{A}\tilde{I}(X;Y \mid Z). \tag{81}$$

We next compute the minimization over A. Note that we can write $\tilde{I}(X;Y \mid Z)$ in an alternative way. Recall from (48) that

$$\begin{aligned}&\tilde{I}(X;Y,Z) =\\ &\log\det(K_{YZ}) - \log\det(K_Z) - \log\det(K_{ME})\end{aligned} \tag{82}$$

where K_{YZ}, K_Z and K_{ME} are covariance matrices. By simplifying the Schur complement of $\det(K_{YZ})$ with $\det(K_Z) = \det(K_X + K_E)$, we get that $\tilde{I}(X;Y \mid Z)$ is given by

$$\begin{aligned}&\log\det(H_MK_XH_M^* + \mathbf{I}_{n_M} -\\ &(H_MK_XH_E^* + A)(H_EK_XH_E^* + \mathbf{I})^{-1}(H_EK_XH_M^* + A^*))\\ &\qquad - \log\det(\mathbf{I}_{n_M} - AA^*).\end{aligned} \tag{83}$$

Proposition 6: Let $\tilde{A}^*$ be a local minima of $\tilde{I}(X;Y \mid Z)$. Then $\tilde{A}^*$ is the solution of an algebraic Ricatti as given in (91).

Proof: Set

$$\begin{aligned}M_1 &= H_MK_XH_M^* + \mathbf{I}_{n_M}\\ M_2 &= H_MK_XH_E^*\\ M_3 &= (H_EK_XH_E^* + \mathbf{I})^{-1}\end{aligned}$$

so that the first term inside the $\log\det$ of (83) can be written as

$$f(A) = M_1 - (A + M_2)M_3(A^* + M_2^*). \tag{84}$$

The gradient of $\log\det f(A)$ with respect to A is

$$\nabla_A\log\det(f(A)) = -f(A)^{-1}(A + M_2)M_3. \tag{85}$$

Using this formula, we compute that

$$\begin{array}{c}\nabla_A \tilde{I}(X;Y|Z)=0 \Longleftrightarrow \\ -f(A)^{-1}(A+M_2)M_3+A(\mathbf{I}-A^*A)^{-1}=0\end{array} \tag{86}$$

with $f(A)$ as in (84). Thus

$$\begin{array}{c}f(A)^{-1}(A+M_2)M_3=A(\mathbf{I}-A^*A)^{-1} \Longleftrightarrow \\ (A+M_2)M_3(\mathbf{I}-A^*A)=f(A)A.\end{array} \tag{87}$$

We now develop $f(A)$ using (84) to get

$$\begin{array}{c}(A+M_2)M_3(\mathbf{I}-A^*A)= \\ M_1A-(A+M_2)M_3(A^*+M_2^*)A \\ \Longleftrightarrow (A+M_2)M_3(\mathbf{I}+M_2^*A)=M_1A.\end{array} \tag{88}$$

This yields the following nonsymmetric algebraic Ricatti equation in A:

$$A[M_3M_2^*]A+A[M_3]+[M_2M_3M_2^*-M_1]A+M_2M_3=\mathbf{0}. \tag{89}$$

We now notice that $\log\det f(A)$ can be alternatively written as

$$\log\det(M_1M_3)+\log\det(M_3^{-1}-(M_2^*+A^*)M_1^{-1}(M_2+A)) \tag{90}$$

where $\log\det(M_1M_3)$ is a constant that disappears in the gradient computation. This yields a gradient expression for A^*, where the roles are exchanged as follows: $A \leftrightarrow A^*, M_2 \leftrightarrow M_2^*, M_1 \leftrightarrow M_3^{-1}$. In turn, we get for A^* the algebraic Ricatti equation

$$\begin{array}{c}A^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_MK_XH_E^*A^*+ \\ A^*[(H_MK_XH_M^*+\mathbf{I})^{-1}]+[-H_EK_XH_E^*-\mathbf{I}+ \\ H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_MK_XH_E^*]A^* \\ +H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}=\mathbf{0}\end{array} \tag{91}$$

with which we will continue the computations. ■

One way of solving an algebraic Riccati [5] of the form

$$0=M_{21}+M_{22}A^*-A^*M_{11}-A^*M_{12}A^* \tag{92}$$

is to look for invariant subspaces of

$$M=\begin{bmatrix}M_{11} & M_{21}\\ M_{12} & M_{22}\end{bmatrix}. \tag{93}$$

Here we have that M is given by

$$\begin{array}{l}M_{11}=-(H_MK_XH_M^*+\mathbf{I})^{-1} \\ M_{12}=-(H_MK_XH_M^*+\mathbf{I})^{-1}H_MK_XH_E^* \\ M_{21}=H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1} \\ M_{22}=-H_EK_XH_E^*-\mathbf{I}+ \\ \qquad H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_MK_XH_E^*.\end{array} \tag{94}$$

Invariant subspaces of M (and consequently explicit solutions to the Ricatti equation) are computed in the following proposition.

Proposition 7: If $n_M+n_E \geq n$, then

$$\tilde{A}^*=(H_EV\ QW)(H_MV\ PW)^{-1} \tag{95}$$

where W is an $(n_M+n_E-n)\times m$ matrix, $0\leq m\leq n_M$, $(P^T\ Q^T)^T$ is orthogonal to $(H_M^*\ H_E^*)$, P,Q of dimension respectively $n_M\times(n_M+n_E-n)$ and $n_E\times(n_M+n_E-n)$, and V is an $n\times(n_M-m)$ matrix, such that

$$\begin{bmatrix}H_MV\\ H_EV\end{bmatrix} \tag{96}$$

is an invariant subspace of M, as defined in (94). If $n_M+n_E<n$, then

$$\tilde{A}^*=(H_E)_{\mathcal{I}}V((H_M)_{\mathcal{I}}V)^{-1} \tag{97}$$

where the subscript $\mathcal{I}$ denotes the matrix obtained by picking the subset $\mathcal{I}$ of linearly independent n_M+n_E columns of the corresponding matrix.

Proof: We start looking for a first invariant subspace of M. Set

$$F=\begin{bmatrix}H_MK_XH_M^*+\mathbf{I}_{n_M} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{n_E}.\end{bmatrix} \tag{98}$$

We have that $F(M+\mathbf{I}_{n_M+n_E})=:B$ is given by

$$\begin{array}{l}B_{11}=H_MK_XH_M^* \\ B_{12}=-H_MK_XH_E^* \\ B_{21}=H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1} \\ B_{22}-H_EK_XH_E^* \\ \quad +H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_MK_XH_E^*.\end{array} \tag{99}$$

It is easy to see that

$$\begin{array}{c}F(M+\mathbf{I})= \\ \begin{bmatrix}-H_M\\ -H_E+H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_M\\ (-K_XH_M^*\ K_XH_E^*)\end{bmatrix}.\end{array} \tag{100}$$

which implies that -1 is an eigenvalue of M if $n_M+n_E>n$. Thus, if $n_M+n_E>n$, a first invariant subspace is given by the eigenspace associated to -1, which is the kernel of $M+\mathbf{I}$, or in other words, the subspace $(P^T\ Q^T)^T$ orthogonal to $(-K_XH_M^*\ K_XH_E^*)$.

For all values of n, n_M, n_E, we can further rewrite M as

$$\begin{array}{rl}F^{-1}\begin{bmatrix}-H_M\\ -H_E+H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_M\\ (-K_XH_M^*\ K_XH_E^*)-\mathbf{I}\end{bmatrix} & (101)\\ =\begin{bmatrix}-(H_MK_XH_M^*+\mathbf{I})^{-1}H_M\\ -H_E+H_EK_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_M\\ (-K_XH_M^*\ K_XH_E^*)-\mathbf{I}\end{bmatrix} & (102)\\ =\begin{bmatrix}-H_M(K_XH_M^*H_M+\mathbf{I})^{-1}\\ -H_E(I-K_XH_M^*(H_MK_XH_M^*+\mathbf{I})^{-1}H_M)\\ (-K_XH_M^*\ K_XH_E^*)-\mathbf{I}\end{bmatrix} & (103)\\ =\begin{bmatrix}-H_M\\ -H_E\end{bmatrix}\cdot & \\ \quad (K_XH_M^*H_M+\mathbf{I})^{-1}(-K_XH_M^*\ K_XH_E^*)-\mathbf{I} & (104)\end{array}$$

using the matrix inversion lemma. Thus, if $n_M+n_E\geq n$, a Jordan basis of M is given by

$$\begin{bmatrix}H_M & P\\ H_E & Q\end{bmatrix} \tag{105}$$

with $(P^T\ Q^T)^T$ orthogonal to $(-H_M^*\ H_E^*)$, since [see (106), shown at the bottom of the page].

In the case where $n_M + n_E < n$, we may assume without loss of generality that the first $n_M + n_E$ columns of

$$\begin{bmatrix} H_M \\ H_E \end{bmatrix} \tag{107}$$

are linearly independent. Denoting them by

$$\begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix} \tag{108}$$

we may write [seee (109), shown at the bottom of the page], where the subscript $\bar{\mathcal{I}}$ denotes the $n - n_M - n_E$ remaining columns of the corresponding matrix. A Jordan basis is thus given by

$$\begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix}. \tag{110}$$

Finally, solutions of the Ricatti equation are given [5], if $n_M + n_E \geq n$, by

$$\tilde{A}^* = (H_E V\ QW)(H_M V\ PW)^{-1} \tag{111}$$

where W is an $(n_M + n_E - n) \times m$ matrix, $0 \leq m \leq n_M$, and V is a $n_M \times (n_M - m)$ matrix, such that

$$\begin{bmatrix} H_M V \\ H_E V \end{bmatrix} \tag{112}$$

is an invariant subspace of M. Note that W can be chosen arbitrary since $(P^T, Q^T)^T$ is the eigenspace associated to -1. The matrix $\tilde{A}^*$ is obtained similarly for the case $n_M + n_E < n$. ■

Corollary 1: In the particular case where both $H_E^* H_E$ and $H_M^* H_M$ are invertible, we get

$$\tilde{A}^* = (H_E(H_M^* H_M)^{-1} H_M^* V\ H_E(H_E^* H_E)^{-1} H_M^* W) \cdot (V, W)^{-1} \tag{113}$$

where V is an $n_M \times (n_M - m)$ matrix, such that

$$\begin{bmatrix} V \\ H_E(H_M^* H_M)^{-1} H_M^* V \end{bmatrix} \tag{114}$$

is an invariant subspace of the matrix M. Furthermore, if $m = n_M$, then $\tilde{A}^* = H_E(H_E^* H_E)^{-1} H_M^*$. Similarly, if $m = 0$, then $\tilde{A}^* = H_E(H_M^* H_M)^{-1} H_M^*$.

Proof: In the case where both $H_E^* H_E$ and $H_M^* H_M$ are invertible, we get the following Jordan basis:

$$\begin{bmatrix} \mathbf{I}_{n_M} & \mathbf{I}_{n_E} \\ H_E(H_M^* H_M)^{-1} H_M^* & H_E(H_E^* H_E)^{-1} H_M^* \end{bmatrix}. \tag{115}$$

■

Proposition 8: Let $\tilde{K}_X$ be an optimal solution to the optimization problem

$$\max_{K_X \succeq \mathbf{0}, Tr(K_X) = P} \min_A \tilde{I}(X; Y \mid Z) \tag{116}$$

where

$$\tilde{A}^* = \begin{cases} (H_E V\ QW)(H_M V\ PW)^{-1} & n_M + n_E \geq n \\ (H_E)_{\mathcal{I}} V((H_M)_{\mathcal{I}} V)^{-1} & n_M + n_E < n \end{cases} \tag{117}$$

is the optimal solution for the minimization over A. Then $\tilde{K}_X$ has rank $r < n$. Furthermore, if $n > n_M$, then $r \leq n_M$.

Proof: We have seen in (76) that $\tilde{I}(X; Y \mid Z)$ can be written

$$\log\det(\mathbf{I} + BK_X) - \log\det(\mathbf{I} + H_E K_X H_E^*) \tag{118}$$

where

$$B := (H_M^* - H_E^* A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E) + H_E^* H_E. \tag{119}$$

That $r \leq n_M$ when $n > n_M$ thus follows from Proposition 2. That $\tilde{K}_X$ is low rank similarly follows from Proposition 2 once we have shown that $B - H_E^* H_E$ is indefinite or semidefinite. We thus prove here that $B - H_E^* H_E$ is low rank.

Let us first rewrite

$$B - H_E^* H_E = (H_M^* - H_E^* A^*)(\mathbf{I} - AA^*)^{-1}(H_M - AH_E). \tag{120}$$

If $n_M + n_E < n$, $B - H_E^* H_E$ is clearly low rank and we are done.

If $n_M + n_E \geq n$, we now show that $B - H_E^* H_E$ is low rank by showing that $(H_M^* - H_E^* A^*)$ is low rank. Indeed, we have that $A^* = (H_E V\ QW)(H_M V\ PW)^{-1}$. Therefore

$$\begin{aligned} &H_M^* - H_E^* A^* \\ &= (H_M^*, -H_E^*) \begin{bmatrix} \mathbf{I} \\ A^* \end{bmatrix} \end{aligned} \tag{121}$$

$$M \begin{bmatrix} H_M & P \\ H_E & Q \end{bmatrix} = \begin{bmatrix} H_M & P \\ H_E & Q \end{bmatrix} \cdot \begin{bmatrix} -(K_X H_M^* H_M + \mathbf{I})^{-1} K_X(-H_M^* H_M + H_E^* H_E) - \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \tag{106}$$

$$\begin{bmatrix} H_M \\ H_E \end{bmatrix} = \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n_M + n_E} & \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix}^{-1} \begin{bmatrix} (H_M)_{\bar{\mathcal{I}}} \\ (H_E)_{\bar{\mathcal{I}}} \end{bmatrix} \end{bmatrix} \tag{109}$$

$$= (H_M^*, -H_E^*) \begin{bmatrix} H_M V & PW \\ H_E V & QW \end{bmatrix} (H_M V, PW)^{-1} \quad (122)$$

which, since $(P^T\ Q^T)^T$ is orthogonal to $(H_M^*\ -H_E^*)$ yields

$$H_M^* - H_E^* A^* = ((H_M^* H_M - H_E^* H_E)V\ \mathbf{0})(H_M V\ PW)^{-1} \quad (123)$$

which, as desired, is low rank. ∎

Proposition 9: Knowing that the rank of K_X is $r < n$, the optimal solution to

$$\min_A \tilde{I}(X; Y \mid Z) \quad (124)$$

is given by

$$\tilde{A}^* = \begin{matrix} (H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ QW)\cdot \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ PW)^{-1} \end{matrix} \quad (125)$$

when $n_M + n_E \geq n$ and

$$\tilde{A}^* = \begin{matrix} (H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ (H_E)_{\mathcal{I}} Q'W)\cdot \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ (H_M)_{\mathcal{I}} Q'W)^{-1} \end{matrix} \quad (126)$$

when $n_M + n_E < n$, where $K_X = U_X U_X^*$, V and W are of dimension respectively $r \times m$ and $n_M \times (n_M - m)$, $m \leq n_M$.

Proof: **Case 1**: $n_M + n_E \geq n$. The Jordan decomposition of M is now given by

$$M \begin{bmatrix} H_M & P \\ H_E & Q \end{bmatrix} = \begin{bmatrix} H_M & P \\ H_E & Q \end{bmatrix} \begin{bmatrix} J & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{bmatrix} \quad (127)$$

where

$$J = (\mathbf{I} + K_X H_M^* H_M)^{-1}(K_X H_M^* - K_X H_E^*) \begin{bmatrix} H_M \\ H_E \end{bmatrix} - \mathbf{I}. \quad (128)$$

Let us now look more carefully at J. We first notice that when K_X is low rank, -1 is an eigenvalue. This is clear since

$$J + \mathbf{I} = (\mathbf{I} + K_X H_M^* H_M)^{-1} K_X (H_M^* H_M - H_E^* H_E) \quad (129)$$

and $\det(K_X) = 0$. Furthermore, since K_X is low rank, it can be factorized as $K_X = U_X^* U_X$ where U_X is a $n \times r$ matrix, if $r < n$ denotes the rank of K_X. Thus

$$J = (\mathbf{I} + K_X H_M^* H_M)^{-1} U_X U_X^* (H_M^* H_M - H_E^* H_E) - \mathbf{I} \quad (130)$$

and clearly $(\mathbf{I} + K_X H_M^* H_M)^{-1} U_X$ is an invariant subspace of J. A Jordan basis is thus given by

$$P' = ((\mathbf{I} + K_X H_M^* H_M)^{-1} U_X\ Q') \quad (131)$$

where Q' is the eigenspace associated to -1. This thus gives us a more precise Jordan basis for M [as defined in (94)], namely

$$\begin{bmatrix} H_M P' & P \\ H_E P' & Q \end{bmatrix} = \begin{bmatrix} H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X & H_M Q' & P \\ H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X & H_E Q' & Q \end{bmatrix} \quad (132)$$

shown at the bottom of the page. In this decomposition, the third block is the eigenspace of -1 of dimension $n_M + n_E - n$ which is always present if $n_M + n_E > n$. The middle block also corresponds to an eigenspace of -1, of dimension $n - r$, this one appearing only when K_X drops rank. The first block is an invariant subspace, corresponding to the r eigenvalues of M that are different from -1. From this Jordan basis of M, if $n_M \geq r$, we have that

$$\tilde{A}^* = \begin{matrix} (H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ QW)\cdot \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ PW)^{-1} \end{matrix} \quad (133)$$

is a solution of the Ricatti equation, where W is any $n_M \times (n_M - m)$ matrix, and V is any $r \times m$ matrix, for $m \leq n_M$.

Case 2: $n_M + n_E < n$. Similarly, the Jordan decomposition of M is given by

$$M \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix} = \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix} J \quad (134)$$

where J is given by

$$\left[\mathbf{I}, \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix}^{-1} \begin{bmatrix} (H_M)_{\bar{\mathcal{I}}} \\ (H_E)_{\bar{\mathcal{I}}} \end{bmatrix} \right] \cdot (\mathbf{I} + K_X H_M^* H_M)^{-1} U_X U_X^* (H_M^*, -H_E^*) \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix} - \mathbf{I}. \quad (135)$$

Since K_X is low rank, then -1 is an eigenvalue of J, and K_X can be factorized as $K_X = U_X U_X^*$ where U_X is a $n \times r$ matrix, if $r < n$ denotes the rank of K_X. Thus, clearly

$$\left[\mathbf{I}, \begin{bmatrix} (H_M)_{\mathcal{I}} \\ (H_E)_{\mathcal{I}} \end{bmatrix}^{-1} \begin{bmatrix} (H_M)_{\bar{\mathcal{I}}} \\ (H_E)_{\bar{\mathcal{I}}} \end{bmatrix} \right] (\mathbf{I} + K_X H_M^* H_M)^{-1} U_X \quad (136)$$

is an invariant subspace of J. Let Q' be the eigenspace associated to -1. Thus, a more precise Jordan basis for M (as defined in (94)) is given by

$$\begin{bmatrix} H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X & (H_M)_{\mathcal{I}} Q' \\ H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X & (H_E)_{\mathcal{I}} Q' \end{bmatrix}. \quad (137)$$

From this Jordan basis of M, we have that

$$\tilde{A}^* = \begin{matrix} (H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ (H_E)_{\mathcal{I}} Q'W) \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\ (H_M)_{\mathcal{I}} Q'W)^{-1} \end{matrix} \quad (138)$$

is a solution of the Ricatti equation, where W is any $n_M \times (n_M m)$ matrix, and V is any $r \times m$ matrix, $m \leq n_M$. ∎

C. The Converse Matches the Achievability

So far, we have solved the optimization problem

$$\min_A \max_{K_X} \tilde{I}(X; Y \mid Z) \quad (139)$$

by computing the optimal $\tilde{A}$ in a closed form expression, and by showing that the optimal $\tilde{K}_X$ is low rank. We are now ready

to conclude the proof, by proving that the optimal $\tilde{A}$ makes the converse match the achievability.

Proposition 10: Let

$$\tilde{A}^* = \begin{matrix}(H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\, QW) \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\, PW)^{-1}\end{matrix} \tag{140}$$

when $n_M + n_E \geq n$ and

$$\tilde{A}^* = \begin{matrix}(H_E(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\, (H_E)_{\mathcal{I}} Q' W) \\ (H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\, (H_M)_{\mathcal{I}} Q' W)^{-1}\end{matrix} \tag{141}$$

when $n_M + n_E < n$ be a solution of the Ricatti equation. Then

$$\begin{aligned}\tilde{I}(X;Y|Z) = &\log\det(\mathbf{I} + H_M K_X H_M^*) \\ &- \log\det(\mathbf{I} + H_E K_X H_E^*).\end{aligned} \tag{142}$$

Furthermore, there exist V, W such that $\mathbf{I} - AA^* \succ \mathbf{0}$.

Proof: Recall from (62) that a way of writing $\tilde{I}(X;Y \mid Z)$ is

$$\begin{aligned}\tilde{I}(X;Y \mid Z) = &\log\det(\mathbf{I} + H_M^* H_M K_X + \\ &(-H_M A + H_E^*)(\mathbf{I} - A^* A)^{-1}(-A^* H_M + H_E) K_X) \\ &- \log\det(\mathbf{I} + H_E K_X H_E^*).\end{aligned} \tag{143}$$

We now show that K_X is in the kernel of $-A^* H_M + H_E$. If $n_M + n_E \geq n$, we have that

$$\begin{aligned}&(H_M(K_X H_M^* H_M + \mathbf{I})^{-1} U_X V\, PW)^{-1} H_M K_X \\ &= (H_M U_X V\, (H_M K_X H_M^* + \mathbf{I}) PW)^{-1} \cdot \\ &\quad (H_M K_X H_M^* + \mathbf{I}) H_M K_X \qquad (144) \\ &= (H_M U_X V\, (H_M K_X H_M^* + \mathbf{I}) PW)^{-1} \cdot \\ &\quad H_M U_X U_X^* (H_M^* H_M K_X + \mathbf{I}) \qquad (145) \\ &= \begin{bmatrix} V^{-1} U_X^* (H_M^* H_M K_X + \mathbf{I}) \\ \mathbf{0} \end{bmatrix} \qquad (146)\end{aligned}$$

so that

$$\begin{aligned}&A^* H_M K_X \\ &= H_E(K_X H_M^* H_M + \mathbf{I})^{-1} K_X (H_M^* H_M K_X + \mathbf{I}) \qquad (147) \\ &= H_E K_X \qquad (148)\end{aligned}$$

so that we get

$$\begin{aligned}\tilde{I}(X;Y|Z) = &\log\det(\mathbf{I} + H_M K_X H_M^*) \\ &- \log\det(\mathbf{I} + H_E K_X H_E^*).\end{aligned} \tag{149}$$

Notice that this computation is similar in the case where $n_M + n_E < n$.

Set $B := K_X H_M^* H_M + \mathbf{I}$. We now have that

$$\begin{aligned}&\mathbf{I} - AA^* \succ \mathbf{0} \\ &\iff \begin{bmatrix} V^* U_X^* B^* H_M^* \\ W^* P^* \end{bmatrix} (H_M B U_X V\, PW) - \\ &\quad \begin{bmatrix} V^* U_X^* B^* H_E^* \\ W^* Q^* \end{bmatrix} (H_E B U_X V\, QW) \succeq \mathbf{0} \qquad (150) \\ &\iff \begin{bmatrix} V^* U_X^* B^* (H_M^* H_M - H_E^* H_E) B U_X V & \mathbf{0} \\ \mathbf{0} & W^* (P^* P - Q^* Q) W \end{bmatrix} \\ &\quad \succ \mathbf{0} \qquad (151)\end{aligned}$$

since

$$V^* U_X^* B^* (H_M^* P - H_E^* Q) W = \mathbf{0} \tag{152}$$

by definition of P and Q. To conclude the proof, notice that when $H_M^* H_M - H_E^* H_E$ is indefinite, there exist V and W such that the above matrix is positive definite. ∎

IV. Conclusion

In this paper, we considered the problem of computing the perfect secrecy capacity of a multiple antenna channel, based on a generalization of the wire-tap channel to a MIMO broadcast wire-tap channel. We proved that for an arbitrary number of transmit/receive antennas, the perfect secrecy capacity is the difference of the two mutual informations, the one of the legitimate user minus the one of the eavesdropper, after a suitable optimization over the transmitters input covariance matrix.

A main assumption in our work is that the transmitter knows the channels to both the legitimate user and the eavesdropper. While this may be a plausible assumption in some scenarios, a more realistic assumption is that the transmitter knows only the statistics of the eavesdropper. An important open problem is to determine the secrecy capacity in this case. Finally, it would also be useful to come up with practical schemes that can guarantee perfect secrecy when the transmitter has only partial CSI of the legitimate user.

Acknowledgment

The authors would like to thank Prof. S. Shamai for pointing out an argument which simplified one of the proofs and the anonymous reviewers for their careful reading of the paper.

References

[1] J. Barros and M. R. D. Rodrigues, "Secrecy capacity of wireless channels," presented at the IEEE Int. Symp. Information Theory, Seattle, WA, Jul. 2006.

[2] M. Bloch, J. Barros, M. R. D. Rodrigues, and S. W. McLaughlin, "Wireless information-theoretic security—Part I: Theoretical aspects," *IEEE Trans. Inf. Theory, Special Issue on Information-Theoretic Security*, vol. 54, no. 6, pp. 2515–2534, Jun. 2008, Special Issue on Information-Theoretic Security.

[3] R. Bustin, R. Liu, H. V. Poor, and S. Shamai, "An MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP J. Wireless Communications and Networking*, no. 3, Mar. 2009, Special Issue on Wireless Physical Layer Security.

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

[5] G. Freiling, "A survey on nonsymmetric Ricatti equations," *Lin. Algebra Appl.*, vol. 351–352, pp. 243–270, 2002.

[6] P. Gopala, L. Lai, and H. El Gamal, "On the secrecy capacity of fading channels," *IEEE Trans Inf. Theory*, vol. 54, no. 11, pp. 5059–5067, Nov. 2008.

[7] A. O. Hero, "Secure space-time communication," *IEEE Trans Inf. Theory*, vol. 49, no. 12, pp. 1–16, Dec. 2003.

[8] A. Khisti, G. Wornell, A. Wiesel, and Y. Eldar, "On the Gaussian MIMO wiretap channel," presented at the Proc. IEEE Int. Symp. Information Theory, Nice, France, 2007.

[9] A. Khisti and G. Wornell, "The MIMOME channel," presented at the 45th Annu. Allerton Conf., Monticello, IL, Sep. 2007.

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

[11] A. Khisti and G. W. Wornell, "Secure transmission with multiple antennas—II: The MIMOME wiretap channel," *IEEE. Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.
[12] S. K. Leung-Yan-Cheong and M. E. Hellman, "The Gaussian wire-tap channel," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 4, pp. 451–456, Jul. 1978.
[13] Z. Li, R. Yates, and W. Trappe, "Secrecy capacity of independent parallel channels," presented at the Allerton Conf., 2006.
[14] Z. Li, W. Trappe, and R. Yates, "Secret communication via multi-antenna transmission," presented at the Conf. Information Sciences and Systems (CISS), Mar. 2007.
[15] Y. Liang and H. V. Poor, "Secure communication over fading channels," presented at the Allerton Conf., 2006.
[16] Y. Liang, H. V. Poor, and S. Shamai (Shitz), "Secure communication over fading channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2470–2492, Jun. 2008, Special Issue on Information Theoretic Security.
[17] R. Liu and H. V. Poor, "Multiple antenna secure broadcast over wireless networks," presented at the 1st Int. Workshop on Information Theory for Sensor Networks, Santa Fe, NM, Jun. 2007.
[18] 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.
[19] T. Liu and S. Shamai (Shitz), "A note on the secrecy capacity of the multi-antenna wiretap channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.
[20] D. G. Luenberger, *Optimization by Vector Space Methods*. Hoboken, NJ: Wiley, 1969.
[21] F. Oggier and B. Hassibi, "The secrecy capacity of the 2×2 MIMO wiretap channel," presented at the 45th Annu. Allerton Conf., Monticello, IL, Sep. 2007.
[22] F. Oggier and B. Hassibi, "The secrecy capacity of the MIMO wiretap channel," presented at the IEEE Int. Symp. Information Theory, Toronto, ON, Canada, Jul. 2008.
[23] P. Parada and R. Blahut, "Secrecy capacity of SIMO and slow fading channels," presented at the IEEE Int. Symp. Information Theory, Adelaide, Australia, 2005.
[24] S. Shafiee and S. Ulukus, "Achievable rates in Gaussian MISO channels with secrecy constraints," presented at the IEEE Int. Symp. Information Theory, Nice, France, Jun. 2007.
[25] S. Shafiee, N. Liu, and S. Ulukus, "Towards the secrecy capacity of the Gaussian MIMO wire-tap channel: The 2-2-1 channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4033–4039, Sep. 2009.
[26] J. A. Thomas, "Feedback can at most double Gaussian multiple access channel capacity," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 711–716, Sep. 1987.
[27] A. D. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, vol. 54, Oct. 1975.

Frédérique Oggier received her degree (Diplôme) in mathematics and computer science in 2000 from the University of Geneva, CH. She then joined the Swiss Federal Institute of Technology, Lausanne (EPFL), where she graduated from the Doctoral School in Communication Systems (2001), and completed her Ph.D. thesis in mathematics (2005). She was a postdoctoral visitor at the California Institute of Technology (CalTech) from 2005 till 2007, and at the Research Center for Information Security (RCIS) in Tokyo, Japan, from 2007 to 2008. She is currently an Assistant Professor at the School of Physical and Mathematical Sciences, Nanyang Technological University (NTU), Singapore. She is a recipient of the Singapore NRF Fellowship.

Her main research interests are in applied algebra to coding problems arising in wireless communications, distributed networked storage as well as information theoretic security.

Babak Hassibi was born in Tehran, Iran, in 1967. He received the B.S. degree from the University of Tehran in 1989, and the M.S. and Ph.D. degrees from Stanford University in 1993 and 1996, respectively, all in electrical engineering.

He has been with the California Institute of Technology since January 2001, where he is currently Professor and Executive Officer of Elecrical Engineering. From October 1996 to October 1998 he was a research associate at the Information Systems Laboratory, Stanford University, and from November 1998 to December 2000 he was a Member of the Technical Staff in the Mathematical Sciences Research Center at Bell Laboratories, Murray Hill, NJ. He has also held short-tem appointments at Ricoh California Research Center, the Indian Institute of Science, and Linkoping University, Sweden. His research interests include wireless communications and networks, robust estimation and control, adaptive signal processing and linear algebra. He is the coauthor of the books (both with A. H. Sayed and T. Kailath) *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories* (SIAM, 1999) and *Linear Estimation* (Prentice Hall, 2000). He is a recipient of an Alborz Foundation Fellowship, the 1999 O. Hugo Schuck best paper award of the American Automatic Control Council (with H. Hindi and S. P. Boyd), the 2002 National Science Foundation Career Award, the 2002 Okawa Foundation Research Grant for Information and Telecommunications, the 2003 David and Lucille Packard Fellowship for Science and Engineering and the 2003 Presidential Early Career Award for Scientists and Engineers (PECASE), and was a participant in the 2004 National Academy of Engineering "Frontiers in Engineering" program.

He has been a Guest Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY special issue on "space-time transmission, reception, coding and signal processing" was an Associate Editor for Communications of the IEEE TRANSACTIONS ON INFORMATION THEORY during 2004–2006, and is currently an Editor for the Journal "Foundations and Trends in Information and Communication".