On the Secrecy Capacity of Rank-Deficient Compound Wiretap Channels

Rafael F. Schaefer and Sergey Loyka

***Abstract*—The secrecy capacity of compound wiretap channels is studied. Using the known lower bounds for non-degraded discrete memoryless channels and establishing the converse, the secrecy capacity is established under the less capable and noisier conditions. The compound capacity is shown to be equal to the worst-case channel capacity when there exists a saddle-point in the mutual information difference.**

Earlier results on the secrecy capacity of compound Gaussian MIMO wiretap channels under the spectral norm constraint are extended to the case of rank-deficient eavesdroppers (which may model a massive MIMO base station) without the degradedness assumption. Its compound secrecy capacity is established in a closed form and the optimal signaling is identified: the compound capacity equals the worst-case channel capacity thus establishing the saddle-point property; the optimal signaling is Gaussian and on the eigenvectors of the legitimate channel and the worst-case eavesdropper is omnidirectional (i.e. isotropic over the sub-space spanned by the active eigenmodes of the legitimate channel).

The case of additive uncertainty in the legitimate channel, in addition to the unknown eavesdropper channel, is also studied and its compound secrecy capacity is established.

I. Introduction

The broadcast nature of wireless communications makes it inherently vulnerable to eavesdropping. The methods of information-theoretic security are instrumental in this regard since they solely use the physical properties of wireless channels to guarantee the secrecy of communications. Information-theoretic security was initiated by Shannon himself and studied later by Wyner, who introduced the now-popular *wiretap channel* modeling the simplest scenario of secure communications with one legitimate transmitter-receiver pair and one wiretapper (eavesdropper) to be kept ignorant of transmitted information [1]. There is presently a growing interest in information-theoretic security, see e.g. [1–3].

Since multi-antenna or multiple-input multiple-output (MIMO) techniques can improve the performance significantly, MIMO architectures are widely accepted as essential for future wireless systems. Their information-theoretic security is currently under active investigation. From the information-theoretic viewpoint, the main performance metric is the secrecy capacity, i.e. the largest achievable rate subject to the reliability and secrecy criteria [1]. The secrecy capacity of the Gaussian MIMO wiretap channel was established in [4–6] under full channel state information (CSI), where it turns out that Gaussian signaling is optimal. Subsequently, the optimal transmit covariance matrix has then been found under the total power constraint for a number of special cases [4, 5, 7–9] but the general case remains illusive.

Due to the dynamic nature of wireless channels as well as the limitations of channel estimation and feedback, wireless systems are subject to channel uncertainty. In addition to this in the secrecy context, it is hardly possible to expect that the eavesdropper (unlike the legitimate receiver) will share its CSI with the transmitter to make the eavesdropping harder, which makes the perfect eavesdropper CSI model more than questionable. A standard approach to this problem is to assume that the exact channel realization is not known; it is only known that it remains fixed during the entire transmission and that it belongs to a known set of channels (uncertainty set), which is known as *compound channel* [10].

The compound wiretap discrete memoryless channel (DMC) with a countably-finite uncertainty set (i.e. finite-state channels) was studied in [11, 12] under the weak and strong secrecy criteria. Its secrecy capacity was established under the degradedness assumption, where all possible realizations of the eavesdropper channel must be degraded with respect to all possible realizations of the legitimate channel. When this condition is not satisfied, only an achievable secrecy rate (i.e. a lower bound to the capacity) was obtained while the secrecy capacity for the general case remains still unknown.

The corresponding compound Gaussian MIMO wiretap channel (WTC) with countably-finite uncertainty sets was studied in [11]. Similarly to the discrete memoryless case, its secrecy capacity was established, again, only under the degradedness assumption. When the channel is not degraded, the secrecy capacity itself remains unknown and only an achievable secrecy rate was obtained. A semi-definite programming algorithm was developed in [13] to characterize numerically an achievable rate for the MISO-WTC with uncertainty when the legitimate receiver has a single antenna. In the case of spherical uncertainty and single-antenna receivers, an analytical solution for an achievable rate was obtained in [14]. It remains unknown how far away is the actual secrecy capacity from these achievable rates.

The secrecy capacity of non-degraded compound MIMO-WTC as well as optimal signaling were established in [15] when there is uncertainty in the eavesdropper channel only, which is bounded by the spectral norm (and hence the uncertainty set is isotropic and uncountably-infinite). The compound capacity was shown to be equal to the worst-case channel capacity so that a code designed for the worst-case channel also works on the whole class of channels. These results were further extended to a broader class of non-isotropic uncertainty sets in [16], where not only the gain but also eigendirections of the eavesdropper channel are bounded. It was shown that

This work was supported by the German Research Foundation (DFG) under Grant WY 151/2-1.

R. F. Schaefer is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA (e-mail: rafaelfs@princeton.edu).

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

978-1-4673-9526-7/15/$31.00 ©2015 IEEE

the worst-case and compound capacities are the same if there exists a maximum element in the uncertainty set.

The studies above make use of uncertainty sets which include full-rank eavesdropper channels and hence may be too conservative if the rank is limited due to e.g. a small number of handset antennas while the transmitter has a large number of antennas (e.g. a massive MIMO base station; this idea was suggested by A. Khisti). In this case, the rank will not exceed the number of handset antennas while the uncertainty model above allows the rank to be equal to a (much larger) number of base station antennas. To model this scenario in the present paper, we explicitly include the rank constraint on the eavesdropper channel. This renders the respective optimization problems non-convex so that the standard tools of convex optimization (e.g. KKT conditions [20] or von Neumann mini-max theorem [19]) cannot be used and new tools have to be developed.

In summary, the main limitations of the known results are that the compound secrecy capacity is only known for degraded channels (DMC) or that the uncertainty set always includes full-rank eavesdroppers (MIMO-WTC). The main difficulty in establishing the secrecy capacity for non-degraded channels seems to be the missing converse. In this paper, we extend the known results in 2 directions. In Part 1, we establish the secrecy capacity of certain discrete memoryless channels (not necessarily Gaussian or degraded); in Part 2, we establish the secrecy capacity of compound Gaussian MIMO channels with rank-constrained eavesdropper and without degradedness assumption. The case of rank-constrained eavesdropper is motivated by the scenario where the transmitter is a base station with a large number of antennas while the receiver/eavesdropper are handsets with a small number of antennas. Under this non-convex constraint (in addition to the convex spectral norm constraint), there is no maximum element in the uncertainty set, yet the saddle-point property is shown to hold and the compound secrecy capacity is established in a closed form along with optimal signaling. Since the respective optimization problems are not convex, neither KKT conditions nor von Neumann mini-max theorem can be used to establish the existence of a saddle-point (a key step in proving the converse - the most difficult step). Hence, we rely on majorization theory and certain novel singular value inequalities to establish the result.

Subsequently, a more general case of two-sided channel uncertainty is studied, where the legitimate channel is also allowed to have (additive) uncertainty. This reflects the assumption that the legitimate receiver will share its CSI with the transmitter, but limitations in feedback link and channel estimation result in channel uncertainty. The corresponding compound secrecy capacity is established and shown to be equal to the secrecy capacity of the worst-case channel in the uncertainty set, so that the saddle-point property still holds. The optimal signaling is still on the eigenmodes of the legitimate channel and the worst-case eavesdropper is omni-directional on the sub-space spanned by the active eigenmodes of the legitimate channel.

II. Compound Wiretap Channel Model

In this section we introduce the compound wiretap DMC model. Building on earlier results in [12, 15], the compound secrecy capacity of non-degraded wiretap channels is established for arbitrary uncertainty sets (not limited to finite-state or countable) under the less capable and noisier conditions. We use the standard model of wiretap DMC extended to the compound setting as explained below.

A. Discrete Memoryless Channels

Let $\mathcal{X}$ and $\mathcal{Y}$, $\mathcal{Z}$ be input and output alphabets and $\mathcal{S}$ be a channel uncertainty set. Unless indicated otherwise, the alphabets are assumed to be discrete and finite in this section. The channels to the legitimate receiver and the eavesdropper (wiretapper) are conditional distributions $W_s : \mathcal{X} \times \mathcal{S} \to \mathcal{P}(\mathcal{Y})$ and $V_s : \mathcal{X} \times \mathcal{S} \to \mathcal{P}(\mathcal{Z})$, respectively, where $s \in \mathcal{S}$ is a channel state and $\mathcal{P}(\cdot)$ is a set of probability distributions. For a fixed state $s \in \mathcal{S}$, input and output sequences $x^n \in \mathcal{X}^n$ and $y^n \in \mathcal{Y}^n$, $z^n \in \mathcal{Z}^n$ of block length n, the discrete memoryless channels are given by $W_s^n(y^n|x^n) = \prod_{i=1}^n W_s(y_i|x_i)$ and $V_s^n(z^n|x^n) = \prod_{i=1}^n V_s(z_i|x_i)$. The channels are assumed to be quasi-static: s is selected at the beginning and is held constant during the entire transmission.

Definition 1. The discrete memoryless *compound wiretap channel* $\mathfrak{W}$ is given by

$$\mathfrak{W} = \big\{(W_s, V_s) : s \in \mathcal{S}\big\}.$$

This includes the widely adopted model of the form $\mathfrak{W} = \{(W_{s_1}, V_{s_2}) : s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}$ with $\mathcal{S}_1 \neq \mathcal{S}_2$ as one can always construct a new set of the form $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2$.

Codebooks and achievable secrecy rate for this channel are defined in the standard way, see e.g. [1, 12, 15]. In particular, an achievable secrecy rate is subject to the reliability and secrecy criteria: the *reliability criterion* implies arbitrary low error probability for the legitimate receiver and the *secrecy criterion* implies arbitrary low information leakage to the eavesdropper (strong secrecy), both as the block length increases to infinity. For the compound channel, this has to be achieved by a single codebook for any channel state in the uncertainty set, i.e. the codebook and legitimate receiver decoding rule must not depend on a particular channel realization s but only on the uncertainty set $\mathcal{S}$ (see e.g. [10]). On the other hand, we make a conservative (and safest from secrecy perspective) assumption that the eavesdropper has perfect channel state information of both channels (to the legitimate receiver and its own).

The wiretap DMC with a countably-finite uncertainty set (i.e. finite-state channels) was studied in [11, 12] and its secrecy capacity has been established under the degradedness assumption. In the general (non-degraded) case, only an achievable secrecy rate is known and the capacity remains unknown (due to the missing converse - the most difficult part). The achievable secrecy rates above were extended to arbitrary uncertainty sets (which are not required to be finite or countable) and to continuous alphabets and compact uncertainty sets in [15].

Theorem 1 ([15]). The compound secrecy capacity C_c of the wiretap DMC $\mathfrak{W}$ is bounded as follows:

$$C_c \geq \sup_{P_X} \big(\inf_{s_1 \in \mathcal{S}_1} I(X;Y_{s_1}) - \sup_{s_2 \in \mathcal{S}_2} I(X;Z_{s_2}) \big) \tag{1}$$

for any uncertainty set $\mathcal{S}$ (not necessarily finite or countable), where I is mutual information, X is the (random) input of distribution P_X while Y_{s_1}, Z_{s_2} are the induced receiver and eavesdropper outputs, respectively, under channel state $s = (s_1, s_2)$. If alphabets are continuous, $\mathcal{S}$ is required to be compact.

However, the secrecy capacity C_c remained unknown otherwise. This gap is filled below. To proceed further, we need the following definitions that establish an ordering of compound wiretap channels.

Definition 2. A compound DMC V_{s_2} is said to be noisier than a compound DMC W_{s_1} if

$$I(U;Y_{s_1}) \geq I(U;Z_{s_2}) \tag{2}$$

for any aggregate channel state $s = (s_1, s_2) \in \mathcal{S}$, any random variable U and any DMC $U \to X$ such that $U \to X \to (Y_{s_1}, Z_{s_2})$ is a Markov chain.

Definition 3. Compound DMC V_{s_2} is said to be (physically) degraded with respect to compound DMC W_{s_1} if $X \to Y_{s_1} \to Z_{s_2}$ is a Markov chain for any channel state $s = (s_1, s_2) \in \mathcal{S}$ and any input X.

These definitions are an extension of the corresponding definition for non-compound (single-state) channels, see e.g. [1, 17]. Similarly to the single-state channels, it can be shown that "degraded" implies "noisier," but the converse is not true, i.e. the latter requirement is weaker than the former (so that there are channels that are "noisier" but not "degraded").

Below, we show that the equality in (1) is achieved under the less noisier condition thus establishing the compound secrecy capacity of this (non-degraded) wiretap DMC.

Theorem 2. If V_{s_2} is noisier than W_{s_1}, the compound secrecy capacity C_c of the discrete memoryless wiretap channel $\mathfrak{W}$ is as follows:

$$C_c = \sup_{P_X} \big(\inf_{s_1 \in \mathcal{S}_1} I(X;Y_{s_1}) - \sup_{s_2 \in \mathcal{S}_2} I(X;Z_{s_2}) \big) \tag{3}$$

for any uncertainty set $\mathcal{S}$ (not necessarily finite or countable).

Proof: To establish the equality under the noisier condition, observe that, by extending the proof of the converse in Theorem 3 of [17] to the compound setting and requiring the encoder to be independent of the actual channel states, it can be shown that any achievable secrecy rate R_s is bounded as follows

$$R_s \leq I(X;Y_{s_1}) - I(X;Z_{s_2}) \tag{4}$$

for a channel state (s_1, s_2), where the input X is induced by the encoder, so that the achievable compound secrecy rate R_c satisfies

$$R_c \leq \inf_s (I(X;Y_{s_1}) - I(X;Z_{s_2})) \tag{5}$$

from which it follows that

$$C_c \leq \sup_{P_X} \inf_s (I(X;Y_{s_1}) - I(X;Z_{s_2})) \tag{6}$$

and thus establishes the equality. ■

Remark 1. Since each degraded channel is also "noisier," the equality in Theorem 1 also holds for degraded channels.

To proceed further, we need the following definitions.

Definition 4. Compound DMC V_{s_2} is said to be less capable than compound DMC W_{s_1} if for every P_X and any channel state $(s_1, s_2) \in \mathcal{S}$

$$I(X;Y_{s_1}) \geq I(X;Z_{s_2}). \tag{7}$$

This definition extends the corresponding definition in [1] to the compound channel setting. Following the same line of analysis as for the single-state channels, it can be shown that the less capable requirement is strictly weaker than the noisier one (i.e. each "noisier" channel is also "less capable" but the converse is not true), and hence strictly weaker than the degraded one.

Definition 5. A compound wiretap channel is said to have a saddle-point if

$$\begin{aligned} \sup_{P_X} \inf_{s \in \mathcal{S}} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \\ = \inf_{s \in \mathcal{S}} \sup_{P_X} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \end{aligned} \tag{8}$$

where $s = (s_1, s_2)$ is the aggregate channel state.

Note that this definition does not impose any operational meaning on the quantities involved. The following Theorem provides such operational meaning.

Theorem 3. If the compound wiretap DMC $\mathfrak{W}$ has a saddle-point and satisfies the less capable condition, then the compound secrecy capacity C_c is the same as the worst-case channel capacity C_w,

$$\begin{aligned} C_c &= \sup_{P_X} \big(\inf_{s_1 \in \mathcal{S}_1} I(X;Y_{s_1}) - \sup_{s_2 \in \mathcal{S}_2} I(X;Z_{s_2}) \big) \\ &= \inf_{s \in \mathcal{S}} \sup_{P_X} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) = C_w. \end{aligned} \tag{9}$$

In particular, the channel has a saddle-point if

1) $\mathcal{S}_1, \mathcal{S}_2$ are compact and convex, and

2) $I(X;Y_{s_1}) - I(X;Z_{s_2})$ is lower semi-continuous and quasi-convex in s, and upper semi-continuous and quasi-concave in P_X.

Proof: Since the legitimate and eavesdropper channel states are independent of each other, it follows that

$$\begin{aligned} \inf_{s_1 \in \mathcal{S}_1} I(X;Y_{s_1}) - \sup_{s_2 \in \mathcal{S}_2} I(X;Z_{s_2}) \\ = \inf_{s \in \mathcal{S}} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \end{aligned} \tag{10}$$

so that the following chain inequality holds

$$\begin{aligned} C_w &= \inf_{s \in \mathcal{S}} \sup_{P_X} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \\ &= \sup_{P_X} \inf_{s \in \mathcal{S}} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \\ &\leq C_c \leq C_w \end{aligned} \tag{11}$$

where first equality holds since, from [1, Corollary 3.5],

$$\sup_{P_X} \big(I(X;Y_{s_1}) - I(X;Z_{s_2}) \big) \tag{12}$$

is the secrecy capacity under channel state (s_1, s_2) and the less capable condition, so that taking $\inf_s$ gives the worst-case capacity; the first inequality is due to Theorem 1 and the last inequality is due to the fact that compound capacity cannot exceed the worst case one (since the compound code has also to work on the worst-case channel). This proves $C_c = C_w$. The last statement follows from von Newmann mini-max theorem and its subsequent generalizations, see e.g. [19, Theorem 9.D]. ■

Remark 2. The importance of this result is due to the fact that a code designed for the worst-case channel also works on the whole class of channels, i.e. is robust (which is not true in general).

Remark 3. The requirement of semi-continuity can be dropped in the case of countably-finite alphabets (since the mutual information is known to be continuous in such settings), but it is essential for countably-infinite or continuous alphabets.

Remark 4. Since "noisier" implies "less capable," Corollary 3 also holds for "noisier" and degraded (physically or stochastically) channels.

B. Continuous Alphabets

While the results above have been established for discrete alphabets, they can also be extended to continuous alphabets using the standard quantization and set partitioning arguments. In particular, Theorem 1 holds for continuous alphabets as well [15] and Theorems 2 and 3 can be similarly extended provided that the uncertainty set $\mathcal{S}$ is compact (the details are omitted due to the page limit).

III. Gaussian MIMO Wiretap Channels

We are now in the position to apply the above results to Gaussian MIMO channels. In this section, we do *not* require the channel to be less capable, noisier or degraded. To this end, let N_T be the number of transmit antennas at the transmitter and $N_{1(2)}$ be the numbers of receive antennas at the legitimate receiver (eavesdropper). The input-output relations for the Gaussian MIMO wiretap channel are given by

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2 \tag{13}$$

where $\mathbf{x} = [x_1, x_2, ..., x_{N_T}]^T \in \mathbb{C}^{N_T \times 1}$ is the transmitted signal, $\mathbf{y}_{1(2)} \in \mathbb{C}^{N_{1(2)} \times 1}$ is the signal at the legitimate receiver (eavesdropper), $\boldsymbol{\xi}_{1(2)} \in \mathbb{C}^{N_{1(2)} \times 1}$ is the circularly-symmetric additive white Gaussian noise at the receiver (eavesdropper) (normalized to unit variance in each dimension), and $\mathbf{H}_{1(2)} \in \mathbb{C}^{N_{1(2)} \times N_T}$ is the matrix of the complex channel gains between each transmit and each receive (eavesdropper) antenna. The channels $\mathbf{H}_{1(2)}$ are assumed to be fixed (constant) during the whole transmission of block length n. We assume an average transmit power constraint $\mathrm{tr}\,\mathbf{R} \le P_T$ where P_T is the total transmit power and $\mathbf{R} = \mathbb{E}\{\mathbf{x}\mathbf{x}^+\}$ is the transmit covariance matrix.

For this channel, the secrecy capacity subject to the total average transmit power constraint is [4–6]

$$C_s = \max_{\mathbf{R}} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} \tag{14}$$

where $\mathbf{W}_i = \mathbf{H}_i^+\mathbf{H}_i$, $i = 1, 2$, and $\max$ is subject to the constraints $\mathbf{R} \ge \mathbf{0}$ and $\mathrm{tr}\,\mathbf{R} \le P_T$.

It is well-known that the problem in (14) is not convex in general and explicit solutions for the optimal transmit covariance are not known for the general case, but only for some special cases (e.g. low-SNR, MISO channels, the full-rank case or weak eavesdropper) [4, 5, 7–9].

Let us now consider a particular compound channel where $\mathbf{H}_1$ is given (known to the transmitter) and $\mathbf{H}_2$ can be any (unknown) subject to the spectral norm constraint

$$\mathbf{H}_2 : |\mathbf{H}_2|_2 = \max_{|\mathbf{x}|=1} |\mathbf{H}_2\mathbf{x}| \le \sqrt{\epsilon} \tag{15}$$

or equivalently

$$\mathbf{W}_2 : |\mathbf{W}_2|_2 = \lambda_1(\mathbf{W}_2) \le \epsilon \tag{16}$$

where $|\mathbf{x}| = \sqrt{\mathbf{x}^+\mathbf{x}}$ is the Euclidean norm of $\mathbf{x}$, $|\mathbf{H}|_2 = \sigma_1(\mathbf{H})$ is the spectral norm of $\mathbf{H}$, i.e. its largest singular value $\sigma_1(\mathbf{H})$; $\lambda_1(\mathbf{W}_2)$ is the largest eigenvalue of $\mathbf{W}_2$. Note that $|\mathbf{H}\mathbf{x}|$ represents the channel (voltage) gain in transmit direction $\mathbf{x}$ so that $|\mathbf{H}|_2$ is the largest channel gain. $|\mathbf{W}|_2$ represents the largest channel power gain. The importance of the spectral norm in the context of MIMO channels has been discussed in [18]. In particular, the set in (15) limits the maximum gain of the eavesdropper channel without putting any constraint on its eigenvectors. This represents the physical scenario where the eavesdropper cannot approach the transmitter beyond a certain minimum (protection) distance (so that the channel gain is bounded due to propagation path loss) being unconstrained otherwise. Note that there is no requirement of degradedness here. The compound secrecy capacity for this class of channels has been established in [16], where it was shown that the saddle-point property holds so that the worst-case and compound capacities are the same and signaling on the worst-case channels works for the entire class of channels. The optimal signaling is Gaussian and on the eigenmodes of $\mathbf{W}_1$ while the optimal power allocation is different from the classical water-filling.

A. Rank-Deficient Eavesdropper

The uncertainty set (15) includes full column-rank eavesdroppers and the worst-case eavesdropper is also full column-rank, which requires the number of eavesdropper antennas to be not less than that of transmit ones, $N_2 \ge N_T$. This may not be the case in certain applications, if e.g. the transmitter is a massive MIMO base station and receivers are handsets with a small number of antennas (due to size/complexity constraint) so that $N_{1,2} \ll N_T$.

In this section, we consider such a scenario by introducing an extra constraint on the rank $r(\mathbf{W}_2) = r(\mathbf{H}_2)$ of the eavesdropper channel $\mathbf{H}_2$, $r(\mathbf{W}_2) \le r_2$ for given $r_2 \le N_T$. This models the fact that $r(\mathbf{W}_2) \le N_2$ so that when the number N_2 of eavesdropper antennas is small, $N_2 < N_T$,

full-rank $\mathbf{W}_2$ is not possible. The resulting eavesdropper uncertainty set is of the form

$$\boldsymbol{\mathcal{S}}_2 = \left\{\mathbf{W}_2 : |\mathbf{W}_2|_2 \le \epsilon,\ r(\mathbf{W}_2) \le r_2\right\} \tag{17}$$

where the 1st inequality reflects the fact that the eavesdropper channel gain is bounded (due to e.g. minimum propagation path loss) and the 2nd one reflects the fact that the rank is bounded due to e.g. small number of eavesdropper antennas. Note that the set $\boldsymbol{\mathcal{S}}_2$ is not convex so that neither KKT conditions nor von Neumann mini-max theorem can be used [19, 20]. The compound secrecy capacity can be explicitly found using the majorization theory and properties of singular values as follows.

Theorem 4. Consider the compound Gaussian MIMO wiretap channel in (13) *with known* $\mathbf{W}_1$ *and unknown* $\mathbf{W}_2$ *belonging to the uncertainty set* $\boldsymbol{\mathcal{S}}_2$ *in* (17)*; assume that* $r(\mathbf{W}_1) = r_1 \le r_2$*. The compound secrecy capacity* C_c *of this channel is the same as the worst-case channel capacity* C_w*,*

$$\begin{aligned} C_c &= \max_{\mathbf{R}} \min_{\mathbf{W}_2} C(\mathbf{R}, \mathbf{W}_2) \\ &= \min_{\mathbf{W}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) = C_w \\ &= C(\mathbf{R}^*, \mathbf{W}_2^*) \\ &= \sum_{i=1}^{r_1} \ln \frac{1 + g_i \lambda_i^*}{1 + \epsilon \lambda_i^*} \end{aligned} \tag{18}$$

where $(\mathbf{R}^*, \mathbf{W}_2^*)$ *is the saddle-point,* max *and* min *are over all admissible* $\mathbf{R}, \mathbf{W}_2$*:* $\mathbf{R}, \mathbf{W}_2 \ge 0$*,* $\operatorname{tr}\mathbf{R} \le P_T$*,* $\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2$*. The optimal signaling is Gaussian and on the eigenmodes of the legitimate channel,*

$$\mathbf{R}^* = \mathbf{U}_1 \boldsymbol{\Lambda}^* \mathbf{U}_1^+, \tag{19}$$

where the columns of unitary matrix $\mathbf{U}_1$ *are the eigenvectors of* $\mathbf{W}_1$*, diagonal matrix* $\boldsymbol{\Lambda} = \operatorname{diag}\{\lambda_i^*\}$ *collects the eigenvalues of* $\mathbf{R}^*$*:*

$$\lambda_i^* = \frac{\epsilon + g_i}{2\epsilon g_i} \left(\sqrt{1 + \frac{4\epsilon g_i}{(\epsilon + g_i)^2} \left(\frac{g_i - \epsilon}{\lambda} - 1 \right)_+} - 1 \right) \tag{20}$$

where $\lambda > 0$ *is found from the total power constraint* $\sum_i \lambda_i^* = P_T$*,* $g_i = \lambda_i(\mathbf{W}_1)$*,* $(x)_+ = \max\{x, 0\}$*. The worst-case eavesdropper is* $\mathbf{W}_2^* = \varepsilon \mathbf{U}_{1a} \mathbf{U}_{1a}^+$*, where the columns of semi-unitary matrix* $\mathbf{U}_{1a}$ *are the eigenvectors of* $\mathbf{W}_1$ *corresponding to strictly positive eigenvalues.*

Proof: Based on the majorization properties of singular values of matrix products and Schur-convexity as applied to $\ln|\mathbf{I} + \mathbf{WR}|$, see [23] for details. ■

Remark 5. Note that the worst-case eavesdropper $\mathbf{W}_2^* = \varepsilon \mathbf{U}_{1a} \mathbf{U}_{1a}^+$ is "isotropic" on the sub-space spanned by the columns of $\mathbf{U}_{1a}$ (but not on the whole space), which is known as "omni-directional" in the antenna literature [21] (i.e. having the same gain in all directions of that sub-space).

Remark 6. Unlike the rank-unconstrained case, there is no dominant channel in the rank-constrained uncertainty set, i.e. $\mathbf{W}_2 \le \mathbf{W}_2^*$ does not hold for all $\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2$, so that the uncertainty set is not "degraded" (with respect to $\mathbf{W}_2^*$ or any other $\mathbf{W}_2$). Since the set $\boldsymbol{\mathcal{S}}_2$ is not convex either, one cannot use von Neumann mini-max Theorem (or its extensions) to infer an existence of saddle-point, which is established in (18) via the singular value inequalities, so that the following inequalities hold for any feasible $\mathbf{R}$ and $\mathbf{W}_2$,

$$C(\mathbf{R}, \mathbf{W}_2^*) \le C_c = C_w = C(\mathbf{R}^*, \mathbf{W}_2^*) \le C(\mathbf{R}^*, \mathbf{W}_2). \tag{21}$$

It can be demonstrated, via the example below, that the saddle-point property does not hold if $r_1 > r_2$.

Example. Consider the following compound channel

$$\mathbf{W}_1 = \mathbf{I},\ |\mathbf{W}_2|_2 \le 1,\ r(\mathbf{W}_2) \le 1 \tag{22}$$

where $m = 2$. It is straightforward to see that

$$C_w = \min_{\mathbf{W}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) = \ln(1 + P_T) \tag{23}$$

$$> \ln(1 + P_T/2) = \max_{\mathbf{R}} \min_{\mathbf{W}_2} C(\mathbf{R}, \mathbf{W}_2) = C_c \tag{24}$$

so that there is a 3 dB loss for any $P_T > 0$ due to the lack of Tx CSI. Note also that $\mathbf{R}^* = P_T \mathbf{I}/2$ in (24) (so that isotropic signaling is optimal) while $\mathbf{R}^* = P_T \mathbf{u}_2 \mathbf{u}_2^+$ in (23) (so that beamforming is optimal), where $\{\mathbf{u}_1, \mathbf{u}_2\}$ are the eigenvectors of $\mathbf{W}_2^* = \mathbf{u}_1 \mathbf{u}_1^+$.

Remark 7. The condition on the ranks $r_1 \le r_2$ is insured if $N_1 \le N_2$ and both channels are of full raw ranks. In particular, this holds if $N_1 = N_2 = 1$.

B. Double-Sided Channel Uncertainty

Here we consider the case where both the legitimate and eavesdropper channels are uncertain. The compound channel model follows the model in (13) where:

$$\boldsymbol{\mathcal{S}}_1 = \left\{\mathbf{H}_1 : \mathbf{H}_1 = \mathbf{H}_0 + \Delta\mathbf{H},\ |\Delta\mathbf{H}|_2 \le \epsilon_1\right\} \tag{25a}$$

$$\boldsymbol{\mathcal{S}}_2 = \left\{\mathbf{H}_2 : |\mathbf{H}_2|_2 \le \epsilon,\ r(\mathbf{H}_2) \le r_2\right\} \tag{25b}$$

where $\mathbf{H}_0$ is the nominal part of $\mathbf{H}_1$ known to the transmitter, and $\Delta\mathbf{H}$ is the uncertain, unknown part. This compound model reflects two important points:

1) The desire of the eavesdropper to be confidential to keep its spying abilities uncompromised, so it does not share its channel with the transmitter and therefore only minimal information about $\mathbf{H}_2$ is available to the latter. Its rank is bounded due to e.g. a small number of antennas.
2) The legitimate receiver, on the other hand, wishes to maximize the rate so it shares its channel with the transmitter. Its channel uncertainty is due to the limitations of the feedback and estimation procedure, which is normally much smaller than that of the eavesdropper (and hence the known nominal part).

Note that while $\boldsymbol{\mathcal{S}}_1$ is convex, $\boldsymbol{\mathcal{S}}_2$ is not so that neither KKT condition nor von Neumann mini-max Theorem are helpful. The secrecy capacity of this compound channel can be characterized via the singular value inequalities as follows. Let

$$C(\mathbf{R}, \mathbf{W}_1, \mathbf{W}_2) = \ln \frac{|\mathbf{I} + \mathbf{W}_1 \mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2 \mathbf{R}|}$$

which depends on the transmit covariance matrix $\mathbf{R}$ and the unknown channels $\mathbf{W}_1 = \mathbf{H}_1^+\mathbf{H}_1$ and $\mathbf{W}_2 = \mathbf{H}_2^+\mathbf{H}_2$ to the legitimate receiver and the eavesdropper respectively.

Theorem 5. Consider the compound Gaussian MIMO wiretap channel in (13) *when* $\mathbf{H}_1$ *and* $\mathbf{H}_2$ *are unknown and belong to the uncertainty sets* $\boldsymbol{\mathcal{S}}_1$ *and* $\boldsymbol{\mathcal{S}}_2$ *in* (25)*. Assume that* $r(\mathbf{H}_0) = r_1 \leq r_2$*. Then, the compound secrecy capacity* C_c *is*

$$\begin{aligned} C_c &= \max_{\mathbf{R}} \min_{\mathbf{W}_1,\mathbf{W}_2} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) \\ &= \min_{\mathbf{W}_1,\mathbf{W}_2} \max_{\mathbf{R}} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) = C_w \\ &= C(\mathbf{W}_{1w}, \mathbf{W}_{2w}, \mathbf{R}^*), \end{aligned} \tag{26}$$

i.e., the worst-case secrecy capacity C_w *is also the (compound) secrecy capacity* C_c *of the class of channels and Gaussian signaling is optimal. The saddle-point property holds,*

$$\begin{aligned} C(\mathbf{W}_{1w}, \mathbf{W}_{2w}, \mathbf{R}) &\leq C_c = C(\mathbf{W}_{1w}, \mathbf{W}_{2w}, \mathbf{R}^*) \\ &\leq C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}^*), \end{aligned} \tag{27}$$

where $(\mathbf{W}_{1w}, \mathbf{W}_{2w}, \mathbf{R}^*)$ *is the saddle-point. The worst-case channel is*

$$\begin{aligned} &\mathbf{W}_{1w} = \mathbf{H}_{1w}^+\mathbf{H}_{1w},\ \mathbf{H}_{1w} = \mathbf{V}_0(\boldsymbol{\Sigma}_0 - \epsilon_1\mathbf{I})_+\mathbf{U}_0^+, \\ &\mathbf{W}_{2w} = \epsilon\mathbf{U}_{0a}\mathbf{U}_{0a}^+,\ \mathbf{H}_{2w} = \mathbf{V}\boldsymbol{\Sigma}_{2w}\mathbf{U}_0^+, \end{aligned} \tag{28}$$

where $\mathbf{U}_0, \mathbf{V}_0$ *are unitary matrices of right and left singular vectors of the nominal channel* $\mathbf{H}_0$ *and* $\boldsymbol{\Sigma}_0$ *is the diagonal matrix of its singular values; semi-unitary matrix* $\mathbf{U}_{0a}$ *collects the columns of* $\mathbf{U}_0$ *corresponding to strictly positive singular values;* $\mathbf{V}$ *is an arbitrary unitary matrix, and*

$$\boldsymbol{\Sigma}_{2w} = diag\{\epsilon, .., \epsilon, 0, .., 0\} \tag{29}$$

is a diagonal matrix with 1st r_1 *diagonal entries being* $\sqrt{\epsilon}$ *and 0 otherwise. The optimal covariance* $\mathbf{R}^*$ *is as in Theorem 4 with the substitution*

$$g_i \rightarrow (\sigma_i(\mathbf{H}_0) - \epsilon_1)_+^2,\ \mathbf{U}_1 \rightarrow \mathbf{U}_0, \tag{30}$$

i.e., the optimal signaling is on the eigenmodes of the worst-case legitimate channel $\mathbf{H}_{1w}$*.*

Proof: Based on the novel matrix singular value inequalities in [22], see [23] for details. ■

Note that this theorem does not require the compound channel to be degraded. Remarkably, the saddle-point property still holds and the worst-case eavesdropper is omni-directional on the sub-space spanned by the active eigenmodes of the nominal legitimate channel; the optimal signaling is similar to that in Theorem 4 (Gaussian signaling is still optimal), with the legitimate channel substituted by its worst-case realization (due to uncertainty). We observe that, as the uncertainty (i.e. ϵ_1 and/or ϵ) increases, fewer and fewer eigenmodes are used until only the strongest one remains active, in which case the beamforming is optimal. From this perspective, beamforming is the most robust strategy (works under largest uncertainty).

While the results above have been established under the total power constraint $tr\mathbf{R} \leq P_T$, using similar reasoning it can be shown that the same result holds under a general power constraint of the form $\mathbf{R} \in \boldsymbol{\mathcal{S}}_{\mathbf{R}}$, where $\boldsymbol{\mathcal{S}}_{\mathbf{R}}$ is a unitary invariant set of positive semi-definite matrices, i.e. $\mathbf{R} \in \boldsymbol{\mathcal{S}}_{\mathbf{R}}$ implies $\mathbf{U}\mathbf{R}\mathbf{U}^+ \in \boldsymbol{\mathcal{S}}_{\mathbf{R}}$ for any unitary $\mathbf{U}$. This constraint limits possible eigenvalues of $\mathbf{R}$ but does not constrain in any way its eigenvectors. Special cases include the total and maximum per-eigenmode power constraints.

REFERENCES

[1] M. Bloch and J. Barros, *Physical-Layer Security: From Information Theory to Security Engineering*. Cambridge University Press, 2011.

[2] E. A. Jorswieck et al, "Guest Editorial: Signal Processing for Wireless Physical Layer Security", *IEEE J. Sel. Areas Commun.*, v. 31, N. 9, pp. 1657-1659, Sep. 2013.

[3] R. F. Schaefer and H. Boche, "Physical Layer Service Integration in Wireless Networks", *IEEE Signal Process. Mag.*, v.31, n. 3, pp. 147-156, May 2014.

[4] 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.

[5] ——, "Secure Transmission With Multiple Antennas–Part II: The MIMOME Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515–5532, Nov. 2010.

[6] F. Oggier and B. Hassibi, "The Secrecy Capacity of the MIMO Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 8, pp. 4961–4972, Aug. 2011.

[7] S. Loyka and C. D. Charalambous, "On Optimal Signaling over Secure MIMO Channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 443–447.

[8] S. Loyka and C. D. Charalambous, "Further Results on Optimal Signaling over Secure MIMO Channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2019–2023.

[9] S. Loyka and C. D. Charalambous, "Rank-Deficient Solutions for Optimal Signaling over Secure MIMO Channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, USA, Jun. 2014, pp. 201–205.

[10] A. Lapidoth, P. Narayan, Reliable Communications Under Channel Uncertainty, IEEE Trans. Information Theory, v. 44, N. 6, pp. 2148-2177, Oct. 1998.

[11] Y. Liang, G. Kramer, H. V. Poor, and S. Shamai (Shitz), "Compound Wiretap Channels," *EURASIP J. Wireless Commun. Netw.*, vol. Article ID 142374, pp. 1–13, 2009.

[12] I. Bjelaković, H. Boche, and J. Sommerfeld, "Secrecy Results for Compound Wiretap Channels," *Probl. Inf. Transmission*, vol. 49, no. 1, pp. 73–98, Mar. 2013.

[13] Q. Li and W.-K. Ma, "Optimal and Robust Transmit Designs for MISO Channel Secrecy by Semidefinite Programming," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3799–3812, Aug. 2011.

[14] J. Li and A. P. Petropulu, "Explicit Solution of Worst-Case Secrecy Rate for MISO Wiretap Channels With Spherical Uncertainty," *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3892–3895, Jul. 2012.

[15] R. F. Schaefer and S. Loyka, "The Secrecy Capacity of a Compound MIMO Gaussian Channel", in *Proc. IEEE Inf. Theory Workshop*, Seville, Spain, Sep. 2013, pp. 104-108.

[16] R. F. Schaefer and S. L. Loyka, "The Compound Secrecy Capacity of a Class of Non-Degraded MIMO Gaussian Channels", in *Proc. 52nd Annu. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, Oct. 2014, pp. 1004-1010.

[17] I. Csiszár and J. Körner, "Broadcast Channels with Confidential Messages," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 339–348, May 1978.

[18] S. Loyka and C. D. Charalambous, "On the Compound Capacity of a Class of MIMO Channels Subject to Normed Uncertainty," *IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2048–2063, Apr. 2012.

[19] E. Zeidler, *Nonlinear Functional Analysis and Its Applications I: Fixed-Point Theorems*. Springer, 1986.

[20] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. Hoboken, New Jersey: Wiley, 2005.

[22] S. Loyka and C. D. Charalambous, "Novel Matrix Singular Value Inequalities and Their Applications to Uncertain MIMO Channels," *IEEE Trans. Inf. Theory*, accepted, Aug. 2015.

[23] R. F. Schaefer and S. Loyka, "The Secrecy Capacity of Compound Gaussian MIMO Wiretap Channels", *IEEE Trans. Inf. Theory*, vol. 61, no. 10, pp. 5535-5552, Oct. 2015.