Fifty-second Annual Allerton Conference
Allerton House, UIUC, Illinois, USA
October 1 - 3, 2014

The Compound Secrecy Capacity of a Class of Non-Degraded MIMO Gaussian Channels

Rafael F. Schaefer

Department of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA
e-mail: rafaelfs@princeton.edu

Sergey L. Loyka

School of Electrical Engineering and Computer Science
University of Ottawa
Ottawa, ON, K1N 6N5, Canada
e-mail: sergey.loyka@ieee.org

***Abstract*—Secrecy capacity of a class of non-degraded compound MIMO Gaussian channels is obtained. Earlier results established for isotropic uncertainty sets are extended to broader class of (non-isotropic) sets, which bound not only the gain but also the eigendirections of the eavesdropper channel. When a maximum element exists in the uncertainty set, a saddle-point exists so that the compound and worst-case channel capacities coincide and signaling on the worst-case channel also works for the whole class of channels. The case of additive uncertainty in the legitimate channel, in addition to the unknown eavesdropper channel of a bounded spectral norm, is also studied. Its compound secrecy capacity and the optimal signaling are established in a closed-form, revealing the saddle-point property. The optimal signaling is Gaussian and on the eigenvectors of the legitimate channel and the worst-case eavesdropper is isotropic. The eigenmode power allocation somewhat resembles the standard water-filling but is not identical to it.**

I. Introduction

Currently, there is a growing interest in information-theoretic security stimulated by wide-spread use of wireless networks [1]. Since multiple-input multiple-output (MIMO) architectures are indispensable for modern wireless systems due to their high spectral efficiency, security aspects of MIMO systems have gained importance as well. The secrecy capacity of the MIMO Gaussian wiretap channel is established in [2–5] under full channel state information (CSI), where it turns out that Gaussian signaling is optimal. The optimal transmit covariance matrix under the total power constraint has then been found for a number of special cases [2, 3, 6, 7] while the general solution remains illusive.

The provision of accurate channel state information to the transmitter (Tx) is a major challenge for wireless communication systems. Along with this, it is hardly possible to expect that the eavesdropper (Ev) will share its CSI with the Tx to make the eavesdropping harder, which makes the perfect Ev CSI model more than questionable. A standard approach to address this problem is via the compound channel model, where the exact CSI is not known to the Tx; it is only known that it remains fixed during the whole transmission duration and that it belongs to a known set of channels (uncertainty set).

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

The discrete memoryless compound wiretap channel with a countably-finite uncertainty set (i.e. finite-state channels) is studied in [8, 9]. The corresponding MIMO Gaussian channel with countably-finite uncertainty sets is analyzed in [8]. Its secrecy capacity is established under the degradedness assumption. When the channel is not degraded, an achievable rate is given while the capacity remains unknown. Interference alignment for the compound MIMO Gaussian wiretap channel is explored in [10]. A MIMO Gaussian wiretap channel where the noiseless Ev channel is arbitrarily varying is considered in [11]. Its achievable secrecy rate is given and the secrecy degrees of freedom are established while its capacity remains unknown.

The secrecy capacity of non-degraded compound MIMO Gaussian channels was established in [12] under the full CSI of the legitimate user (Rx) and an Ev uncertainty set subject to the spectral norm constraint (which is uncountably-infinite). This represents the scenario where perfect feedback exists for the Rx channel while the Ev channel is not known, but is known to have a bounded channel gain (due to e.g. propagation path loss). This automatically implies only a minimal Ev CSI at the Tx, which reflects well the natural eavesdropper desire to be confidential and its lack of cooperation. The compound capacity was shown to be equal to the worst-case channel capacity thus establishing a saddle point. The optimal signaling is Gaussian and on the eigenmodes of the worst-case channel, while the eigenmode power allocation somewhat resembles the classical water-filling (WF) but is not identical to it [12].

In this paper, we study the case where the Ev uncertainty set is non-isotropic (not only the gain but also the eigendirections are bounded). This is motivated by the fact that there may be some physical (non-isotropic) constraints in the propagation environment which limit possible eigendirections of the eavesdropper channel. Its compound capacity is characterized via maximum and maximal elements of the uncertainty set. In particular, the existence of a maximum element is sufficient for the saddle point to exist so that the compound capacity equals the worst-case channel capacity and the signaling on the latter is optimal for the whole class of channels. When

978-1-4799-8009-3/14/$31.00 ©2014 IEEE

the maximum element does not exist, the optimal solution is shown to be on a part of the boundary of the uncertainty set so that optimization over a reduced set of maximal elements is sufficient for the overall optimality (see Propositions 1-3 for details).

Next, we study the case of double-sided channel uncertainty, when both the legitimate and eavesdropper channels are not known precisely at the Tx. The Rx channel is allowed to have (additive) uncertainty, which represents channel estimation and feedback link limitations, while the Ev channel has a bounded spectral norm (due to the propagation path loss). No degradedness assumption is made. This is motivated by the Rx channel estimation inaccuracy, in addition to the lack of knowledge of eavesdropper channel. Under this conditions, the saddle point is shown to exist, so that the compound and worst-case capacities are the same and the signaling on the worst-case channel in the uncertainty set is optimal for the whole class of channels. The optimal signaling directions are the eigenmodes of the worst-case Rx channel and the optimal power allocation is somewhat similar but not identical to the WF (see Theorem 2 for details).

II. MIMO Gaussian Wiretap Channel

Let N_T and $N_{1(2)}$ be the numbers of antennas at the transmitter and legitimate receiver (eavesdropper). The MIMO Gaussian wiretap channel is then given by

$$\mathbf{y}_1 = \mathbf{H}_1\mathbf{x} + \boldsymbol{\xi}_1, \quad \mathbf{y}_2 = \mathbf{H}_2\mathbf{x} + \boldsymbol{\xi}_2 \tag{1}$$

where $\mathbf{x} = [x_1, x_2, ..., x_{N_T}]^T \in \mathbb{C}^{N_T\times 1}$ is the Tx signal, $\mathbf{y}_{1(2)} \in \mathbb{C}^{N_{1(2)}\times 1}$ is the signal at the Rx (Ev), $\boldsymbol{\xi}_{1(2)} \in \mathbb{C}^{N_{1(2)}\times 1}$ 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 channel matrix. The channels $\mathbf{H}_{1(2)}$ are assumed to be fixed (constant) during the whole transmission of block length n. We assume an average 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.

The secrecy is insured by requiring the information leakage to the eavesdropper to vanish, which implies that its bit error probability P_b approaches $1/2$ as $n \to \infty$ (and thus codeword error probability approaches 1) and the speed of convergence depends on the secrecy criterion adopted. In particular, it can be shown that

$$P_b = 1/2 - o(1) \text{ under weak secrecy,}$$
$$P_b = 1/2 - o\left(1/\sqrt{n}\right) \text{ under strong secrecy,}$$

so that $P_b \to 1/2$ in any case, but the speed of convergence can be arbitrarily slow under weak secrecy, while it is at least as $1/\sqrt{n}$ under strong secrecy. Using the recent result in [9] on exponential convergence of information leakage to zero, it can be further shown that

$$P_b = 1/2 - \mathcal{O}(e^{-an})$$

i.e. exponentially fast in that scenario, where $a > 0$. This provides an operational meaning for the secrecy criteria.

For the channel in (1), the secrecy capacity subject to the total average transmit power constraint is [2–5]

$$C_s = \max_{\mathbf{R}\ge\mathbf{0}} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \mathbf{W}_2\mathbf{R}|} \qquad \text{s.t.} \quad \mathrm{tr}\,\mathbf{R} \le P_T \tag{2}$$

where $\mathbf{W}_i = \mathbf{H}_i^+\mathbf{H}_i$, $i = 1, 2$, and $^+$ is Hermitian conjugation.

The problem in (2) 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, or for the full-rank case) [2–6].

III. Eavesdropper Channel Uncertainty

Let us consider a compound channel with single-sided uncertainty, where $\mathbf{H}_1$ in (1) is known to the transmitter and $\mathbf{H}_2$ can be any (unknown to the Tx) subject to the spectral norm constraint

$$\begin{aligned}\boldsymbol{\mathcal{S}}_2 =& \left\{\mathbf{H}_2 : |\mathbf{H}_2|_2 = \max_{|\mathbf{x}|=1} |\mathbf{H}_2\mathbf{x}| \le \sqrt{\epsilon}\right\} \\ =& \left\{\mathbf{W}_2 : |\mathbf{W}_2|_2 = \lambda_1(\mathbf{W}_2) \le \epsilon\right\}\end{aligned} \tag{3}$$

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 $\mathbf{W}_2$. Thus, the set $\boldsymbol{\mathcal{S}}_2$ includes *all* $\mathbf{W}_2$ that are less than or equal to $\epsilon\mathbf{I}$.

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 set in (3) limits the maximum gain of the eavesdropper channel without putting any constraint on its eigenvectors. This represents the physical scenario where the Ev 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. The secrecy requirement must hold for all possible Ev channels in $\boldsymbol{\mathcal{S}}_2$ simultaneously. Throughout the paper, full CSI at the eavesdropper is assumed (the safest assumption from the secrecy perspective).

The compound secrecy capacity of this channel has been established in [12], which is summarized below.

Theorem 1 ([12]). Consider the compound MIMO Gaussian wiretap channel in (1) *with known* $\mathbf{W}_1$ *and unknown* $\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2$ *as in* (3)*. Its compound secrecy capacity* C_c *equals to the worst-case channel capacity* C_w*,*

$$C_c = \max_{\mathbf{R}} \min_{\mathbf{W}_2} C(\mathbf{R}, \mathbf{W}_2) = \min_{\mathbf{W}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) = C_w$$

where

$$C(\mathbf{R}, \mathbf{W}_2) = \ln|\mathbf{I} + \mathbf{W}_1\mathbf{R}| - \ln|\mathbf{I} + \mathbf{W}_2\mathbf{R}|, \tag{4}$$

$\max$ *and* $\min$ *are over all admissible* $\mathbf{R}, \mathbf{W}_2$*:* $\mathbf{R}, \mathbf{W}_2 \ge \mathbf{0}$*,* $\mathrm{tr}\,\mathbf{R} \le P_T$*,* $\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2$*, and*

$$C_w = C^*(\epsilon) = \max_{\mathrm{tr}\,\mathbf{R}\le P_T} C(\mathbf{R}, \epsilon\mathbf{I}) \tag{5}$$

is the secure capacity for the isotropic Ev $\mathbf{W}_{2w} = \epsilon\mathbf{I}$*, which is the worst-case one in* $\boldsymbol{\mathcal{S}}_2$*. The optimal signaling is on the eigenmodes of the legitimate channel,*

$$\mathbf{R}^* = \mathbf{U}_1\boldsymbol{\Lambda}^*\mathbf{U}_1^+, \tag{6}$$

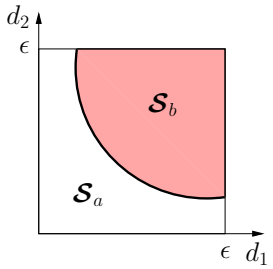

Fig. 1. An example of two uncertainty sets when $\mathbf{W}_2 = \mathrm{diag}\{d_1, d_2\} \geq 0$. The (whole) set $\boldsymbol{\mathcal{S}}_a$ corresponds to the uncertainty set given in (3), while the shaded set $\boldsymbol{\mathcal{S}}_b$ corresponds to (8).

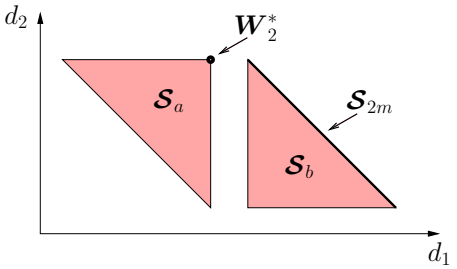

Fig. 2. An example of two uncertainty sets $\boldsymbol{\mathcal{S}}_a$ and $\boldsymbol{\mathcal{S}}_b$ when $m = 2$ and $\mathbf{W}_2 = \mathrm{diag}\{d_1, d_2\} \geq 0$. $\boldsymbol{\mathcal{S}}_a$ has a (unique) maximum element $\mathbf{W}_2^*$ (dark dot) while $\boldsymbol{\mathcal{S}}_b$ does not, but only a set of maximal elements (dark line) $\boldsymbol{\mathcal{S}}_{2m}$.

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

$$\lambda_i^* = \frac{\epsilon + g_i}{2\epsilon g_i} z_i, z_i = \sqrt{1 + \frac{4\epsilon g_i}{(\epsilon + g_i)^2} \left(\frac{g_i - \epsilon}{\lambda} - 1 \right)_+} - 1 \quad (7)$$

and $\lambda > 0$ is found from the total power constraint $\sum_i \lambda_i^* = P_T$, $g_i = \lambda_i(\mathbf{W}_1)$, $(x)_+ = \max\{x, 0\}$. ■

This result shows that the secrecy capacity of the worst-case channel is also the (compound) secrecy capacity of the class of channels (achievable by a single code on the whole class).

Moreover, it follows that the isotropic eavesdropper is the worst-case one under a bounded channel gain for *any* $\mathbf{W}_1$. This is also appealing from the channel feedback perspective: it is hardly possible to expect that the eavesdropper will share its channel with the transmitter to make eavesdropping harder, so only minimal information can be expected by the transmitter about the eavesdropper channel.

A. Broader Class of Compound MIMO Channels

Theorem 1 can be further extended to a broader class of compound MIMO channels. To this end, let us generalize the uncertainty set $\boldsymbol{\mathcal{S}}_2$ for the eavesdropper channel as follows

$$\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2 \rightarrow \mathbf{W}_2 \leq \epsilon \mathbf{I} \in \boldsymbol{\mathcal{S}}_2, \quad (8)$$

i.e., all its members are less than or equal to $\epsilon\mathbf{I}$. Unlike (3), it may include *not all* such $\mathbf{W}_2$; it is not required to be convex, compact etc. Fig. 1 illustrates the difference between the uncertainty sets defined in (3) and (8) for diagonal $\mathbf{W}_2$.

Proposition 1. Consider the compound MIMO Gaussian wiretap channel in (1) *when $\mathbf{W}_1$ is known and unknown $\mathbf{W}_2$ belongs to the uncertainty set $\boldsymbol{\mathcal{S}}_2$ in* (8)*. Its compound secrecy capacity is $C_c = C^*(\epsilon)$, i.e., as in Theorem 1.*

Proof: Observe that the compound secrecy capacity of this channel is not smaller than that in Theorem 1, since the uncertainty set here is included in the uncertainty set of Theorem 1 (which includes *all* $\mathbf{W}_2 \leq \epsilon\mathbf{I}$, since it is equivalent to $\lambda_1(\mathbf{W}_2) \leq \epsilon$). On the other hand, setting $\mathbf{W}_2 = \epsilon\mathbf{I}$ demonstrates that the lower bound is achieved by this worst-case channel. Since the compound capacity does not exceed the worst-case one, the desired result follows. ■

We remark that the set $\boldsymbol{\mathcal{S}}_2$ is not necessarily isotropic, convex or compact (as in (3)), nor it has some other "nice" properties, except that $\epsilon\mathbf{I}$ is its dominant (maximum) element, and that Theorem 1 is a special case. This demonstrates the importance of the isotropic eavesdropper for compound MIMO wiretap channels under minimum Ev CSI at the Tx, even when the uncertainty set is not isotropic.

To generalize these results further, we will need the following definitions.

Definition 1. Let $\boldsymbol{\mathcal{S}}_2$ be an uncertainty set of $\mathbf{W}_2$. $\mathbf{W}_2^*$ is its (unique) maximum element if $\mathbf{W}_2^* \in \boldsymbol{\mathcal{S}}_2$ and $\forall \mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2 \rightarrow \mathbf{W}_2 \leq \mathbf{W}_2^*$.

Definition 2. $\mathbf{W}_{2m}$ is a maximal element of $\boldsymbol{\mathcal{S}}_2$ if $\mathbf{W}_2, \mathbf{W}_{2m} \in \boldsymbol{\mathcal{S}}_2$ and $\mathbf{W}_2 \geq \mathbf{W}_{2m} \rightarrow \mathbf{W}_2 = \mathbf{W}_{2m}$ (i.e. the only element in $\boldsymbol{\mathcal{S}}_2$ greater or equal to $\mathbf{W}_{2m}$ is $\mathbf{W}_{2m}$ itself).

Note that Definition 2 is due to the fact that not any two positive semi-definite matrices can be compared (i.e. it can be that neither $\mathbf{W}_1 \geq \mathbf{W}_2$ nor $\mathbf{W}_1 < \mathbf{W}_2$ is true, unlike the scalar case), so that a maximum element may not exist. While maximum element, if it exists, is unique, there may be many maximal elements in a set (see e.g. [14] for more details). Fig. 2 illustrates these definitions for the case of diagonal $\mathbf{W}_2$ and $m = 2$. We are now able to generalize Proposition 1.

Proposition 2. Consider the compound MIMO Gaussian wiretap channel in (1) *when $\mathbf{W}_1$ is known and unknown $\mathbf{W}_2$ belongs to an arbitrary uncertainty set $\boldsymbol{\mathcal{S}}_2$, whose maximum element is $\mathbf{W}_2^*$. The saddle-point property holds, so that the compound secrecy capacity equals to the worst-case secrecy capacity:*

$$\begin{aligned} C_c &= \max_{\mathbf{R}} \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2) \\ &= \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) \\ &= \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2^*) \end{aligned} \quad (9)$$

where the worst-case channel is $\mathbf{W}_2^*$, and the transmission on this channel is optimal for the whole class of channels in $\boldsymbol{\mathcal{S}}_2$.

Proof: Observe that

$$C(\mathbf{R}, \mathbf{W}_2) \geq C(\mathbf{R}, \mathbf{W}_2^*) \ \forall \mathbf{R}, \mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2$$

which is due to the fact that $|\mathbf{I} + \mathbf{WR}|$ is monotonically increasing in $\mathbf{W}$ [13] for any (positive semi-definite) $\mathbf{R}$, so that, by using $\max\min$ and $\min\max$ on both sides,

$$\begin{aligned}\max_{\mathbf{R}} \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2) &= \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2^*) \\ &= \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2)\end{aligned}$$

which proves the desired result. ∎

This proposition says, in effect, that the saddle-point property holds and, thus, the compound secrecy capacity equals to the worst-case one, if a maximum element of the uncertainty set $\boldsymbol{\mathcal{S}}_2$ exists and the rest of its structure is irrelevant. Under this condition, a code designed for the worst-case channel works for the whole class of compound channels.

When the uncertainty set does not have a maximum element, its worst-case secrecy capacity can be characterized using maximal elements as follows.

Proposition 3. Consider the compound MIMO Gaussian channel in (1) *when* $\mathbf{W}_1$ *is known and unknown* $\mathbf{W}_2$ *belongs to a bounded and closed uncertainty set* $\boldsymbol{\mathcal{S}}_2$*, which does not have a maximum element. Then,*

$$\min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2) = \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_{2m}} C(\mathbf{R}, \mathbf{W}_2) \ \forall \mathbf{R} \tag{10}$$

where $\boldsymbol{\mathcal{S}}_{2m}$ *is the set of all maximal elements* $\mathbf{W}_{2m}$ *of* $\boldsymbol{\mathcal{S}}_2$*, and the worst-case secrecy capacity is*

$$C_w = \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) = \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_{2m}} \max_{\mathbf{R}} C(\mathbf{R}, \mathbf{W}_2) \tag{11}$$

i.e. minimizing over the whole uncertainty set $\boldsymbol{\mathcal{S}}_2$ *is equivalent to minimizing over (normally much smaller) set of its maximal elements.*

Proof: The proof is relegated to Appendix A. ∎

We remark that Proposition 3 effectively reduces the dimensionality of the related optimization problem: if the original problem in (11) is D-dimensional, the reduced one (on the right hand side) is at most $(D-1)$-dimensional, since $\boldsymbol{\mathcal{S}}_{2m}$ is on the boundary of $\boldsymbol{\mathcal{S}}_2$ (this can be proved by contradiction).

The last two propositions demonstrate the key role of the maximum element in the uncertainty set: if it exists, a saddle-point exists, so it is a sufficient condition. It can be shown, via examples, that the absence of a maximum element may or may not result in the absence of a saddle-point, so there is no necessary condition here.

IV. Double-Sided Channel Uncertainty

Here we consider the case where both the legitimate and eavesdropper channels are not known precisely at the Tx, i.e. double-sided channel uncertainty. The compound channel model follows the model in (1) where:

$$\boldsymbol{\mathcal{S}}_1 = \left\{\mathbf{H}_1 : \mathbf{H}_1 = \mathbf{H}_0 + \Delta\mathbf{H}, \ |\Delta\mathbf{H}|_2 \leq \epsilon_1\right\} \tag{12a}$$

$$\boldsymbol{\mathcal{S}}_2 = \left\{\mathbf{W}_2 : |\mathbf{W}_2|_2 \leq \epsilon\right\} \tag{12b}$$

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; $|\Delta\mathbf{H}|_2 = \sigma_1(\Delta\mathbf{H})$ is the spectral norm of $\Delta\mathbf{H}$, i.e. the largest singular value $\sigma_1(\Delta\mathbf{H})$. The uncertainty of $\mathbf{W}_2$ follows the same model as in (3). This compound model reflects two important points:

First, 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. Second, 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).

Let us define

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

which depends on the transmit covariance matrix $\mathbf{R}$ and the unknown channels $\mathbf{W}_1 = \mathbf{H}_1^+\mathbf{H}_1$ and $\mathbf{W}_2 = \mathbf{H}_2^+\mathbf{H}_2$. The secrecy capacity of the compound channel in (12) can now be characterized as follows.

Theorem 2. Consider the compound MIMO Gaussian wiretap channel in (1) *when fixed* $\mathbf{W}_1$ *and* $\mathbf{W}_2$ *are unknown at the Tx but known to belong to the uncertainty sets* $\boldsymbol{\mathcal{S}}_1$ *and* $\boldsymbol{\mathcal{S}}_2$ *in* (12)*. Then, the compound secrecy capacity* C_c *is*

$$\begin{aligned}C_c &= \max_{\mathbf{R}} \min_{\mathbf{W}_1, \mathbf{W}_2} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) \\ &= \min_{\mathbf{W}_1, \mathbf{W}_2} \max_{\mathbf{R}} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) = C_w \\ &= C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*),\end{aligned} \tag{13}$$

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

$$\begin{aligned}C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}) &\leq C_c = C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*) \\ &\leq C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}^*),\end{aligned} \tag{14}$$

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

$$\begin{aligned}&\mathbf{W}_{1w} = \mathbf{H}_{1w}^+\mathbf{H}_{1w}, \ \mathbf{H}_{1w} = \mathbf{V}_0(\boldsymbol{\Sigma}_0 - \epsilon_1\mathbf{I})_+\mathbf{U}_0^+, \\ &\mathbf{W}_{2w} = \epsilon\mathbf{I},\end{aligned} \tag{15}$$

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. The optimal covariance* $\mathbf{R}^*$ *is as in Theorem 1 with the substitution*

$$g_i \to (\sigma_i(\mathbf{H}_0) - \epsilon_1)_+^2, \ \mathbf{U}_1 \to \mathbf{U}_0, \tag{16}$$

i.e., the optimal signaling is on the eigenmodes of the worst nominal channel $\mathbf{H}_{1w}$ and isotropic eavesdropper.

Proof: The proof can be found in Appendix B. ∎

Remarkably, the saddle-point property holds and the isotropic eavesdropper (of the maximum gain) is still the worst-case one, even under the legitimate channel uncertainty, and the optimal signaling is almost the same as in Theorem 1, with the legitimate channel substituted by its degraded (due to uncertainty) version $\mathbf{H}_{1w}$. 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.

The game-theoretic interpretation of the inequalities in (14) is the same as for the single-sided uncertainty: $\{\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*\}$ is a saddle-point in the matrix game between the transmitter on one side and the eavesdropper and nature on the other; neither can deviate from the optimal strategy without incurring a penalty provided that the other player follows the strategy.

V. Weak vs. Strong Secrecy

The results above have been established under the strong secrecy condition. It was demonstrated in [16, 17] that, for regular (non-compound or known channels), strong and weak secrecy capacities are the same. That result, however, does not immediately apply to the compound setting here. Nevertheless, it can be shown that the weak C_c^{weak} and strong C_c^{strong} secrecy compound capacities are the same,

$$C_c^{weak} = C_c^{strong} \tag{17}$$

if the saddle-point property holds under strong secrecy, i.e. $C_w = C_c^{strong}$. Indeed, under the saddle point property,

$$C_w = C_c^{strong} \leq C_c^{weak} \leq C_w \tag{18}$$

from which (17) follows, where we have used the fact that the worst-case capacity is the same under the strong and weak secrecies, and that the compound strong secrecy capacity is not larger than the weak one. In particular, the results in Theorems 4, 5 and Proposition 5 also hold under weak secrecy, so that one can go from weak to strong secrecy for free in the compound settings as well under the saddle-point property.

In fact, the chain argument in (18) has the following implications:

- the saddle point under strong secrecy ($C_w = C_c^{strong}$) implies a saddle point under weak secrecy ($C_w = C_c^{weak}$),
- no saddle point under weak secrecy ($C_w > C_c^{weak}$) implies no saddle point under strong secrecy ($C_w > C_c^{strong}$).

VI. Conclusion

In this paper, the compound wiretap channel has been studied. The (strong) secrecy capacity of a class of non-degraded compound MIMO Gaussian wiretap channels has been established under the spectral norm constraint on the eavesdropper channel. The channel is not required to be degraded. The optimal signaling as well as the secrecy capacity are given in a closed form. The saddle-point property has been shown to hold, so that the compound capacity equals to the worst-case one and signaling on the worst-case channel achieves the compound capacity. Isotropic eavesdropper is the worst-case one and signaling on the eigenmodes of the legitimate channel is optimal. The results are extended to non-isotropic uncertainty sets. It is shown that the existence of a maximum element in the uncertainty set is sufficient for a saddle-point to exist, so that compound capacity equals to the worst-case one and signaling on the worst-case channel achieves the capacity of the whole class of channels. Finally, these results are extended to include the legitimate channel uncertainty.

Acknowledgment

The authors are grateful to P. Mitran for insightful discussions and suggestions.

Appendix

A. Proof of Proposition 3

The following lemma is instrumental.

Lemma 1. Let $\mathbf{W}_1, \mathbf{W}_2, ...$ be a bounded and increasing sequence of positive semi-definite matrices, i.e.

$$\mathbf{0} \leq \mathbf{W}_1 \leq \mathbf{W}_2 \leq .. \leq \mathbf{W}_i \leq ... \leq a\mathbf{I} \tag{19}$$

where $0 < a < \infty$ is a positive constant. This sequence converges.

Proof: Consider the following sequence of (non-negative) scalars $\alpha_i = \mathbf{x}^+\mathbf{W}_i\mathbf{x}$, where $\mathbf{x}$ is a vector of appropriate size; for convenience, we take $|\mathbf{x}| = 1$. Since $\{\mathbf{W}_i\}$ is an increasing and bounded sequence, so is $\{\alpha_i\}$,

$$0 \leq \alpha_1 \leq \alpha_2 \leq .. \leq \alpha_i \leq ... \leq a \tag{20}$$

and therefore it converges to some non-negative number $b(\mathbf{x}) = \lim_{i\to\infty} \alpha_i \leq a$. Hence, for any $\epsilon > 0$, there is such $n(\epsilon, \mathbf{x})$ that $b(\mathbf{x}) - \alpha_i < \epsilon \ \forall i > n(\epsilon, \mathbf{x}), \mathbf{x}$. Since this is true for any $\mathbf{x}$, take $n(\epsilon) = \max_{\mathbf{x}} n(\epsilon, \mathbf{x})$ and observe that $|b(\mathbf{x}) - \alpha_i| < \epsilon \ \forall i > n(\epsilon)$ and all $\mathbf{x}$. It follows that $\{\alpha_i\}$ is a Cauchy sequence, i.e. $|\alpha_j - \alpha_i| < \epsilon \ \forall i, j > n(\epsilon)$ and all $\mathbf{x}$, i.e.

$$\mathbf{x}^+(\mathbf{W}_j - \mathbf{W}_i)\mathbf{x} < \epsilon \ \forall \mathbf{x}$$

from which it follows that $\lambda_1(\mathbf{W}_j - \mathbf{W}_i) < \epsilon$ and thus $\|\mathbf{W}_j - \mathbf{W}_i\| \to 0$ in any norm (since all norms are equivalent [13]), i.e. $\{\mathbf{W}_i\}$ is a Cauchy sequence and thus converges [18, 19], $\mathbf{W}_i \to \mathbf{W} \leq a\mathbf{I}$. Taking Frobenius norm, one obtains element-wise convergence of this matrix sequence. ∎

Note that this result generalizes to matrices the well-known fact that any scalar increasing and bounded sequence converges.

To proceed further, observe from the definition of $\boldsymbol{\mathcal{S}}_{2m}$ that

$$\min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2) \leq \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_{2m}} C(\mathbf{R}, \mathbf{W}_2). \tag{21}$$

We prove the equality by contradiction. Assume that

$$\min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2) < \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_{2m}} C(\mathbf{R}, \mathbf{W}_2) \tag{22}$$

and let $\mathbf{W}_2^* = \arg\min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_2} C(\mathbf{R}, \mathbf{W}_2)$ be a minimizer over $\boldsymbol{\mathcal{S}}_2$. Then, $\mathbf{W}_2^* \notin \boldsymbol{\mathcal{S}}_{2m}$ (due to the strict inequality) so that there exists $\mathbf{W}_{21} \in \boldsymbol{\mathcal{S}}_2$ such that $\mathbf{W}_{21} \geq \mathbf{W}_2^*$ (otherwise $\mathbf{W}_2^*$ were in $\boldsymbol{\mathcal{S}}_{2m}$), $\mathbf{W}_{21} \neq \mathbf{W}_2^*$, and $C(\mathbf{R}, \mathbf{W}_{21}) \leq C(\mathbf{R}, \mathbf{W}_2^*)$. If $\mathbf{W}_{21} \in \boldsymbol{\mathcal{S}}_{2m}$, we have a contradiction:

$$\begin{aligned} C(\mathbf{R}, \mathbf{W}_{21}) &\leq C(\mathbf{R}, \mathbf{W}_2^*) \\ &< \min_{\mathbf{W}_2 \in \boldsymbol{\mathcal{S}}_{2m}} C(\mathbf{R}, \mathbf{W}_2) \\ &\leq C(\mathbf{R}, \mathbf{W}_{21}). \end{aligned} \tag{23}$$

Assume further that $\mathbf{W}_{21} \notin \boldsymbol{\mathcal{S}}_{2m}$ so that there exists such $\mathbf{W}_{22} \in \boldsymbol{\mathcal{S}}_2$ that $\mathbf{W}_{22} \geq \mathbf{W}_{21}$, $\mathbf{W}_{22} \neq \mathbf{W}_{21}$, and the process is repeated. In this way, we construct a non-decreasing, bounded sequence $\{\mathbf{W}_2^*, \mathbf{W}_{21}, ..., \mathbf{W}_{2i}, ...\}$, which either terminates in a finite number of steps (when some $\mathbf{W}_{2k} \in \boldsymbol{\mathcal{S}}_{2m}$ so we cannot find a greater one) or it continues indefinitely. In the first case, we have a contradiction and thus the assertion is proved.

In the second case, we claim that the sequence will converge to some $\mathbf{W} \in \boldsymbol{\mathcal{S}}_{2m}$. To see this, first observe that this sequence will converge to some $\mathbf{W} \in \boldsymbol{\mathcal{S}}_2$ (due to Lemma 1, since $\boldsymbol{\mathcal{S}}_2$ is bounded and closed and thus compact and the sequence is increasing and bounded; the boundedness can be understood in any norm, since all matrix norms are equivalent). Thus, we have to prove that $\mathbf{W} \in \boldsymbol{\mathcal{S}}_{2m}$. To see this, first observe that $\mathbf{W} \geq \mathbf{W}_{2i}\ \forall i$ (since the sequence is increasing). If $\mathbf{W} \notin \boldsymbol{\mathcal{S}}_{2m}$, then there exists $\mathbf{W}^* \in \boldsymbol{\mathcal{S}}_2$ such that $\mathbf{W}^* \geq \mathbf{W} \geq \mathbf{W}_{21}$ so it can be taken as a part of the constructed sequence and thus $\mathbf{W}$ cannot be its limit - a contradiction. Therefore, $\mathbf{W} \in \boldsymbol{\mathcal{S}}_{2m}$, as claimed. This, however, results in a contradiction to (22) so that (10) holds. To see (11), take $\max_{\mathbf{R}}$ in (21)-(23) and apply the same argument.

B. Proof of Theorem 2

First, we observe that

$$C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) \geq C(\mathbf{W}_1, \epsilon\mathbf{I}, \mathbf{R})\ \forall \mathbf{R}, \mathbf{W}_1, \tag{24}$$

since $\mathbf{W}_2 \leq \epsilon\mathbf{I}$ (which follows from $|\mathbf{W}_2|_2 \leq \epsilon$) and $|\mathbf{I}+\mathbf{W}\mathbf{R}|$ is monotonically increasing in $\mathbf{W}$ for any (positive semi-definite) $\mathbf{R}$. The lower bound is achieved by $\mathbf{W}_2 = \epsilon\mathbf{I}$. Therefore,

$$\min_{\mathbf{W}_2} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}) = C(\mathbf{W}_1, \epsilon\mathbf{I}, \mathbf{R})\ \forall \mathbf{R}, \mathbf{W}_1, \tag{25}$$

and also

$$\begin{aligned} C_w &= \min_{\mathbf{W}_1} \max_{\mathbf{R}} C(\mathbf{W}_1, \epsilon\mathbf{I}, \mathbf{R}) \\ &= \min_{\mathbf{W}_1} \max_{\mathbf{R}} \ln \frac{|\mathbf{I} + \mathbf{W}_1\mathbf{R}|}{|\mathbf{I} + \epsilon\boldsymbol{\Lambda}|} \\ &\overset{(a)}{=} \min_{\mathbf{W}_1} \max_{\mathbf{R}} \sum_i \ln \frac{1 + \lambda_i(\mathbf{W}_1)\lambda_i(\mathbf{R})}{1 + \epsilon\lambda_i(\mathbf{R})} \\ &\overset{(b)}{=} \max_{\{\lambda_i\}} \sum_i \ln \frac{1 + (\sigma_i(\mathbf{H}_0) - \epsilon_1)_+^2\lambda_i}{1 + \epsilon\lambda_i} \\ &= C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*) \end{aligned} \tag{26}$$

where (a) follows from the inequality

$$|\mathbf{I} + \mathbf{W}_1\mathbf{R}| \leq \prod_i (1 + \lambda_i(\mathbf{W}_1)\lambda_i(\mathbf{R})) \tag{27}$$

and the equality is achieved when $\mathbf{W}_1, \mathbf{R}$ have the same eigenvectors; (b) follows from the inequality $\sigma_i(\mathbf{H}_1) \geq (\sigma_i(\mathbf{H}_0) - \sigma_1(\Delta\mathbf{H}))_+$ (see e.g. [13]) and $\lambda_i(\mathbf{W}_1) = \sigma_i^2(\mathbf{H}_1)$ where the equality is achieved by $\mathbf{H}_{1w}$.

We further observe that the saddle-point property in (13) is equivalent to (see e.g. [15])

$$C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}) \overset{(a)}{\leq} C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*) \overset{(b)}{\leq} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}^*) \tag{28}$$

and we prove these inequalities below thus establishing (13).

Note that (a) follows from (26) (since $\mathbf{R}^*$ is the optimal covariance for $\mathbf{W}_1 = \mathbf{W}_{1w}, \mathbf{W}_2 = \epsilon\mathbf{I}$). To prove (b), we need the following technical lemma, which is an extension of well-known singular value inequalities for a sum and a product of two matrices (see e.g. [13]):

Lemma 2. Let $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$ *be* $n \times m$ *and* $m \times m$ *matrices, and let the right singular vectors of* $\mathbf{A}$ *be the same as the left singular vectors of* $\mathbf{C}$ *so that their singular value decompositions (SVD) are* $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}_a\mathbf{V}^+$ *and* $\mathbf{C} = \mathbf{V}\boldsymbol{\Sigma}_c\mathbf{W}^+$, *where* $\mathbf{U}, \mathbf{V}, \mathbf{W}$ *are unitary and* $\boldsymbol{\Sigma}_a = \mathrm{diag}\{\sigma_{ai}\}, \boldsymbol{\Sigma}_c = \mathrm{diag}\{\sigma_{ci}\}$ *are "diagonal" matrices of singular values of* $\mathbf{A}$ *and* $\mathbf{C}$. *Assume that* $\{\sigma_{ai}\}$ *and* $\{\sigma_{ci}\}$ *are in decreasing order. Then,*

$$\sigma_i((\mathbf{A} + \mathbf{B})\mathbf{C}) \geq (\sigma_i(\mathbf{A}) - \sigma_1(\mathbf{B}))_+\sigma_i(\mathbf{C}) \tag{29}$$

where $\sigma_i((\mathbf{A} + \mathbf{B})\mathbf{C})$ *are also in decreasing order. The equality is achieved by* $\mathbf{B} = -\mathbf{U}\boldsymbol{\Sigma}_b\mathbf{V}^+$, *where* $\boldsymbol{\Sigma}_b = \mathrm{diag}\{\min(\sigma_i(\mathbf{A}), \varepsilon)\}$. ■

Using this lemma, one obtains:

$$\begin{aligned} C_w &= C(\mathbf{W}_{1w}, \epsilon\mathbf{I}, \mathbf{R}^*) \\ &\overset{(a)}{=} \sum_i \ln \frac{1 + (\sigma_i(\mathbf{H}_0) - \epsilon_1)_+^2\lambda_i^*}{1 + \epsilon\lambda_i^*} \\ &\overset{(b)}{\leq} \sum_i \ln \frac{1 + \sigma_i^2(\mathbf{H}_1\mathbf{R}^{*1/2})}{1 + \epsilon\lambda_i^*} \\ &= C(\mathbf{W}_1, \epsilon\mathbf{I}, \mathbf{R}^*) \\ &\overset{(c)}{\leq} C(\mathbf{W}_1, \mathbf{W}_2, \mathbf{R}^*) \end{aligned} \tag{30}$$

where (a) follows from (26), (b) follows from Lemma 2 applied to $\mathbf{A} = \mathbf{H}_0, \mathbf{B} = \Delta\mathbf{H}, \mathbf{C} = \mathbf{R}^{*1/2}$ (and observing, from (7), that the singular values of $\mathbf{H}_0$ and $\mathbf{R}^{*1/2}$ are ordered likewise), where we have used $\lambda_i(\mathbf{R}) = \sigma_i^2(\mathbf{R}^{1/2})$, and (c) follows from (24). This establishes (28) and thus (13).

REFERENCES

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

[2] 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.

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

[4] 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.

[5] T. Liu and S. Shamai (Shitz), "A Note on the Secrecy Capacity of the Multiple-Antenna Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 6, pp. 2547–2553, Jun. 2009.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] A. Khisti, "Interference Alignment for the Multiantenna Compound Wiretap Channel," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2976–2993, May 2011.

[11] X. He and A. Yener, "MIMO Wiretap Channels with Arbitrarily Varying Eavesdropper Channel States", http://arxiv.org/abs/1007.4801.

[12] 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.

[13] F. Zhang, *Matrix Theory: Basic Results and Techniques*, 2nd ed. Springer, 2011.

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

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

[16] I. Csiszár, "Almost Independence and Secrecy Capacity," *Probl. Pered. Inform.*, vol. 32, no. 1, pp. 48–57, 1996.

[17] U. M. Maurer and S. Wolf, "Information-Theoretic Key Agreement: From Weak to Strong Secrecy for Free," in *EUROCRYPT 2000, Lecture Notes in Computer Science*. Springer-Verlag, May 2000, vol. 1807, pp. 351–368.

[18] W. Miller, *Symmetry Groups and Their Applications*. Academic Press Inc, 1972.

[19] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2003.