

On the Capacity of a Class of MIMO Channels Subject to Normed Uncertainty

Sergey Loyka

School of Information Technology and Engineering
University of Ottawa
161 Louis Pasteur, Ottawa, Canada, K1N 6N5
Email: sergey.loyka@ieee.org

Charalambos D. Charalambous

University of Cyprus
75 Kallipoleos Avenue, P.O. Box 20537
Nicosia, 1678, Cyprus
Email: chadcha@ucy.ac.cy

Abstract— The compound capacity of uncertain MIMO channels is considered, when the channel is modeled by a class described by an induced norm constraint. Within this framework, two types of classes are investigated, namely, additive and multiplicative uncertainties subject to a spectral norm constraint, using partial channel state information at the transmitter side. The compound capacity is defined as a maxmin of the mutual information, corresponding to the capacity of the class, in which the minimization is done over the class of channels while the maximization is done over the transmit covariance. Closed form solutions for the compound capacity of the classes are obtained while several properties related to transmit and received eigenvectors are presented. It is also shown that capacity of the class of channels is equal to the worst-case channel capacity, while establishing a saddle-point property. Additionally, explicit closed-form solutions are given for the capacity-achieving Tx covariance matrix and the worst-case channel uncertainty. The effect of uncertainty is shown to be equivalent to an SNR loss which is proportional to the size of the uncertainty of the channel matrix measured by the spectral norm.

I. INTRODUCTION

Since the pioneering work of Foschini [1] and Telatar [2], multiple-input multiple-output (MIMO) wireless systems have attracted a significant attention due to a promise of high spectral efficiency. A lot of research activities have been concentrated on both the information-theoretic limits (e.g. capacity) and practical transmission schemes to approach those limits [3]. Performance of such transmission schemes depends heavily on channel state information (CSI) available at the receiver (Rx)/transmitter (Tx) or both. In addition, channel knowledge, either complete or partial, and its accuracy affect also the channel capacity [8]. In wireless systems, imperfect CSI may be due to time-varying nature of the channel as well as due to limitations of the estimation technique and feedback channel [9]. A number of models have been developed to take into account such effects [8][9], but most of them were not specifically tailored for MIMO systems. A concise review of recent results on the impact of imperfect CSI on MIMO system performance can be found in [12]. The models of channel with imperfect CSI can be classified into statistical (when the true channel is considered to be random with given mean and covariance) and deterministic (when the channel is deterministic (fixed), but only known to belong to a certain class, i.e. compound channel [8]) [12]. Deterministic channel uncertainty models have been used to evaluate the performance of zero-forcing precoding and detection techniques [10], of orthogonal space-time block codes [11] and

to evaluate the compound capacity of uncertain rank-one (Ricean) MIMO channels [12].

In this paper, we introduce uncertainty models for the channel matrix based on an induced (spectral) norm constraint, since for practical purposes the spectral norm of the channel matrix has the interpretation of the maximum input/output transfer gain. Following this framework, we develop generic (e.g. any rank) multiplicative and additive channel uncertainty models with the spectral norm constraint specifically tailored for MIMO capacity analysis. In our approach, we follow the ideas pioneered in [4][5]; a comprehensive review of recent research activities can be found in [8]. We consider a MIMO channel matrix \mathbf{H} as consisting of two parts: the nominal channel \mathbf{H}_0 , which is known at both the Tx and Rx, and the uncertainty (perturbation) $\Delta\mathbf{H}$, which is not known, so that $\mathbf{H} = \mathbf{H}_0 + \Delta\mathbf{H} \in S_H$, where S_H is a known (bounded) class of channels; both \mathbf{H}_0 and $\Delta\mathbf{H}$ are assumed to be fixed during the transmission interval, while $\Delta\mathbf{H}$ is unknown at both the Tx and Rx ends. Unlike previous work, the size of $\Delta\mathbf{H}$ is measured by the spectral norm, which is an induced norm. This is fundamentally different from the trace norm considered in [12], which is equivalent to a Frobenius norm that is not an induced norm and thus does not possess a number of important properties of the latter (see Section II-A for details). Furthermore, while the analysis in [12] was limited to a rank-one nominal channel, we consider arbitrary-rank nominal channels. Following [4][5][8], we consider two types of problems, 1) the capacity of the worst case-channel in the class and, 2) the capacity of the class of channels also known as the compound channel capacity [8] that is achievable by a single code for any channel in the class. The former is formulated as a minimax optimization problem while the later is formulated as a maxmin optimization. Via explicit evaluation of these two capacities, we demonstrate that, for both the multiplicative and additive uncertainty models under the spectral norm constraint, the saddle point property holds: the capacity of the class equals to the worst-case channel capacity. Additionally, we give explicit expressions for the optimum (capacity-achieving) covariance of the Tx signal and the worst-case uncertainty of the channel, both of which depend only on the nominal channel and the “size” of the uncertainty set. The effect of uncertainty with bounded spectral norm is shown to be equivalent to an SNR loss (in terms of the total SNR for the multiplicative uncertainty and the per-channel SNR for the additive uncertainty). Similar uncertainty model has been also

considered in [13], where, however, the uncertainty was limited to the singular values of the channel only (no uncertainty in singular vectors), which is very unlikely from the physical perspective. Our model is more general as it does not limit uncertainty to the singular values only (this formulation of the problem makes it more challenging and requires new tools to solve).

II. UNCERTAIN CHANNEL MODEL

Let us consider the following baseband discrete-time MIMO channel model,

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\xi} \quad (1)$$

where $\mathbf{x} = [x_1, x_2, \dots, x_m]^T \in C^{m,1}$ and $\mathbf{y} = [y_1, y_2, \dots, y_n]^T \in C^{n,1}$ are the vectors representing the Tx and Rx symbols respectively, “ T ” denotes transposition, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_m] \in C^{n,m}$ is the $n \times m$ matrix of the complex channel gains between each Tx and each Rx antenna, where \mathbf{h}_i denotes i -th column of \mathbf{H} , n and m are the numbers of Rx and Tx antennas respectively, without loss of generality, we further assume $n \geq m$, $\boldsymbol{\xi}$ is the vector of circularly-symmetric additive white Gaussian noise (AWGN), which is independent and identically distributed (i.i.d.) in each receiver¹. The channel is assumed to be quasistatic (i.e. constant for a sufficiently long period of time so that the infinite horizon information theory assumption holds) and frequency-flat, with partial channel state information (CSI) at the Rx and Tx ends, as described below.

A. Multiplicative Uncertainty

We consider first the multiplicative channel uncertainty model,

$$\mathbf{H} = (\mathbf{I} + \mathbf{E})\mathbf{H}_0 \quad (2)$$

where \mathbf{H}_0 is the nominal channel (without uncertainty) known at the Rx and, possibly, at the Tx, and $\mathbf{E} \in C^{n,n}$ is the multiplicative uncertainty. We assume \mathbf{E} , and hence \mathbf{H} , belong to a limited uncertainty set (i.e. limited measurement error²),

$$\mathbf{H} \in S_H = \left\{ \mathbf{H} : \|\mathbf{E}\|_2 \leq \varepsilon \right\}, \quad \|\mathbf{E}\|_2 = \max_{|\mathbf{x}| \neq 0} \frac{|\mathbf{E}\mathbf{x}|}{|\mathbf{x}|} = \sigma_1(\mathbf{E}) \leq \varepsilon < 1 \quad (3)$$

where S_H is the multiplicative uncertainty set, $\sigma_i(\mathbf{E})$ is i -th singular value of \mathbf{E} , \mathbf{u}_i is the i -th left singular vector, \mathbf{v}_i is the i -th right singular vector, the singular values are ordered as $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$, and $|\mathbf{x}|^2 = \mathbf{x}^+ \mathbf{x} = \sum_i |x_i|^2$ is the vector length squared, and $+$ denotes Hermitian conjugation. Using (2)(3), the norm of the channel uncertainty $\Delta\mathbf{H} = \mathbf{H} - \mathbf{H}_0$ can be bounded as $\|\Delta\mathbf{H}\|_2 \leq \varepsilon \|\mathbf{H}_0\|_2$.

The use of the spectral norm for uncertainty has a number of advantages:

- As a vector induced norm, it lower bounds any matrix norm (including Frobenius) [6] and thus gives the largest uncertainty set S_H . It is also an indicator of the strongest

eigenmode of the uncertainty.

- From the input (\mathbf{x}) - output ($\mathbf{E}\mathbf{x}$) point of view, \mathbf{v}_1 is the highest gain input direction while \mathbf{u}_1 is the highest gain observing direction.

- It is the only unitary-invariant vector-induced norm [6]. Unitary invariance represents the fact that errors in measurements are statistically equal in all directions, i.e. no preferred direction or bias.

- For a unit energy Rx signal coming from the nominal channel ($|\mathbf{H}_0\mathbf{x}|^2 = 1$), the spectral norm of \mathbf{E} limits the energy in the uncertain part of the Rx signal ($|\mathbf{E}\mathbf{H}_0\mathbf{x}|^2 \leq \sigma_1^2(\mathbf{E})$) so that the power ratio of uncertain and certain portions of the Rx signal is upper bounded by $\sigma_1^2(\mathbf{E})$.

- If the uncertain part of the Rx signal is modeled as AWGN, then the corresponding degradation in SNR can be easily evaluated and the new degraded SNR is $\gamma_0 / (1 + \gamma_0 \sigma_1^2(\mathbf{E}))$, where γ_0 is the SNR in the nominal channel. In this model, the effect of uncertainty can be neglected if $\sigma_1(\mathbf{E}) \ll 1 / \sqrt{\gamma_0}$, which immediately gives a rough idea as to how good the channel estimation should be: Under the model in (2), (3), $1 - \sigma_1(\mathbf{E}) \leq \sigma_i(\mathbf{H}) / \sigma_i(\mathbf{H}_0) \leq 1 + \sigma_1(\mathbf{E})$, where $\sigma_i(\mathbf{H})$ is i -th singular value of \mathbf{H} . Combining the last two inequalities, $|1 - \sigma_i(\mathbf{H}) / \sigma_i(\mathbf{H}_0)| \ll 1 / \sqrt{\gamma_0}$, i.e. the normalized uncertainty in the channel singular values should be much less than $1 / \sqrt{\text{SNR}}$ for its effect to be negligible.

B. Additive Uncertainty

In this model, the nominal channel \mathbf{H}_0 experiences an additive perturbation $\Delta\mathbf{H}$,

$$\mathbf{H} = \mathbf{H}_0 + \Delta\mathbf{H} \quad (4)$$

where we also assume that $\Delta\mathbf{H}$, and hence \mathbf{H} , belong to a limited uncertainty set,

$$\mathbf{H} \in S_H = \left\{ \mathbf{H} : \|\mathbf{H} - \mathbf{H}_0\|_2 = \sigma_1(\Delta\mathbf{H}) \leq \varepsilon \right\}, \quad (5)$$

where S_H is the additive uncertainty set. Similarly to multiplicative uncertainty, the spectral norm bound in (5) can be interpreted in terms of the uncertain signal power, the corresponding degradation in SNR and the condition for its negligible effect can be evaluated.

We note that the additive and multiplicative uncertainty models are related, albeit in a non-symmetric way: for any multiplicative uncertainty \mathbf{E} , there exists an equivalent additive uncertainty $\Delta\mathbf{H} = \mathbf{E}\mathbf{H}_0$; the converse is not always true: for given additive uncertainty $\Delta\mathbf{H}$ there exists an equivalent multiplicative uncertainty \mathbf{E} only if the system of linear equations $\Delta\mathbf{H} = \mathbf{E}\mathbf{H}_0$ has a solution.

III. CAPACITY OF MIMO CHANNELS

For fixed \mathbf{H} and given covariance of the Tx vector signal $\mathbf{R} = \mathbf{x}\mathbf{x}^+$, where \mathbf{x} denotes expectation of \mathbf{x} , the mutual information between \mathbf{x} and \mathbf{y} when \mathbf{x} is Gaussian (i.e. capacity-achieving) is given by the celebrated Foschini-Telatar formula,

¹ the case of unequal noise power per Rx can also be considered within the present framework.

² If the limit $\varepsilon < 1$ is not set, the worst-case capacity becomes zero.

$$I = \log \left| \mathbf{I} + \frac{\gamma}{m} \mathbf{H} \mathbf{R} \mathbf{H}^+ \right| = \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} \sigma_i^2(\mathbf{H} \mathbf{R}^{1/2}) \right) \quad (6)$$

$$= I(\mathbf{H}, \mathbf{R})$$

where γ is the SNR, $\sigma_i(\mathbf{H} \mathbf{R}^{1/2})$ are singular values of $\mathbf{H} \mathbf{R}^{1/2}$, and, due to the total Tx power constraint, $tr(\mathbf{R}) \leq m$, where the equality provides the maximum mutual information. If no CSI is available at the Tx end, the popular choice is $\mathbf{R} = \mathbf{I}$. When CSI is available at the Tx end, the capacity can be found as,

$$C(\mathbf{H}) = \max_{tr(\mathbf{R}) \leq m} I(\mathbf{H}, \mathbf{R}) \quad (7)$$

The maximum in (7) has a well-known water-filling solution [1],

$$C(\mathbf{H}) = \max_{\sum_i \lambda_i(\mathbf{R}) \leq m} \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} \sigma_i^2(\mathbf{H}) \lambda_i(\mathbf{R}) \right)$$

$$= \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} \sigma_i^2(\mathbf{H}) \lambda_i(\mathbf{R}_b) \right) \quad (8)$$

$$= \sum_{i_+} \log \left(\frac{\mu \gamma}{m} \sigma_i^2(\mathbf{H}) \right)$$

where $\lambda_i(\mathbf{R}_b) = [\mu - m / (\gamma \sigma_i^2(\mathbf{H}))]_+$ are the eigenvalues of the best (maximizing) covariance matrix \mathbf{R}_b , $[x]_+ = x$ if $x > 0$ and 0 otherwise, $i_+ = \{i : \lambda_i(\mathbf{R}) > 0\}$ is the index of active modes, and the constant μ is found from the total power constraint, $\sum_{i_+} \lambda_i(\mathbf{R}_b) = m$.

IV. CAPACITY OF A CLASS OF MIMO CHANNELS WITH UNCERTAINTY

Below we consider the capacity of channels with uncertainty models in (2)-(5). Following the framework developed in [4][5], we distinguish between the capacity $C_{\min \max}$ of the worst-case channel and the capacity $C_{\max \min}$ of the class of channels (compound capacity) in S_H ,

$$C_{\min \max} = \min_{\mathbf{H} \in S_H} \max_{tr(\mathbf{R}) \leq m} I(\mathbf{H}, \mathbf{R}) \quad (9)$$

$$C_{\max \min} = \max_{tr(\mathbf{R}) \leq m} \min_{\mathbf{H} \in S_H} I(\mathbf{H}, \mathbf{R}) \quad (10)$$

Note that the capacity of the class $C_{\max \min}$ (compound capacity) is achievable using a single code for any channel in the class [4][5], meaning there exists a code so that the decoding error tends to zero uniformly over the class of channels, while the worst-case capacity requires a code specifically tailored to the worst-case channel (which may perform worse on other channels in the class). In general, the following inequality holds [4],

$$C_{\min \max} \geq C_{\max \min} \quad (11)$$

i.e. the capacity of the class is never higher than the worst-case channel capacity. Below we obtain stronger results for the uncertainty models in (2), (4).

A. Multiplicative uncertainty

The capacities in (9) and (10) can be characterized in a simple way, using the following Lemma.

Lemma 1: For given \mathbf{R} , the worst-case mutual information in (6) for the class of channels in (2), (3) is

$$I_{\min}(\mathbf{H}_0, \mathbf{R}) = \min_{\mathbf{H} \in S_H} I(\mathbf{H}, \mathbf{R})$$

$$= \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} (1-\varepsilon)^2 \sigma_i^2(\mathbf{H}_0 \mathbf{R}^{1/2}) \right) \quad (12)$$

i.e. the multiplicative uncertainty in (2), (3) results in relative SNR loss $(1-\varepsilon)^2$ compared to the nominal channel case.

Proof: using the following singular value inequalities [6],

$$\sigma_n(\mathbf{A}) \sigma_i(\mathbf{B}) \leq \sigma_i(\mathbf{A} \mathbf{B}), \quad (13)$$

$$\sigma_i(\mathbf{B}) - \sigma_1(\mathbf{C}) \leq \sigma_i(\mathbf{B} + \mathbf{C}), \quad i = 1 \dots \min(m, n)$$

where $\mathbf{A} \in C^{n,n}$, $\mathbf{B}, \mathbf{C} \in C^{n,m}$, together with (2), (3) results in the lower bound on the mutual information,

$$I(\mathbf{H}, \mathbf{R}) \geq \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} (1-\varepsilon)^2 \sigma_i^2(\mathbf{H}_0 \mathbf{R}^{1/2}) \right) \quad (14)$$

from which (12) follows. Q.E.D.

We note that the worst-case channel perturbation \mathbf{E}_{\min} , which achieves I_{\min} , is of the form,

$$\mathbf{E}_{\min} = -\mathbf{U}'_0 \Sigma_\varepsilon \mathbf{U}'_0{}^+ \quad (15)$$

where $\Sigma_\varepsilon = \text{diag}(\varepsilon \dots \varepsilon)$, and the columns of \mathbf{U}'_0 are the left singular vectors of $\mathbf{H}_0 = \mathbf{H}_0 \mathbf{R}^{1/2}$, which can be found from its singular value decomposition (SVD), $\mathbf{H}_0 = \mathbf{U}'_0 \Sigma'_0 \mathbf{V}'_0{}^+$.

Theorem 1: The capacity of the class of channels in (3) is

$$C_{\max \min} = \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} (1-\varepsilon)^2 \sigma_i^2(\mathbf{H}_0) \lambda_i(\mathbf{R}_b) \right)$$

$$= \sum_{i_+} \log \left(\frac{\mu \gamma}{m} (1-\varepsilon)^2 \sigma_i^2(\mathbf{H}_0) \right) \quad (16)$$

where $\mathbf{R}_b = \mathbf{R}_b(\mathbf{H}_0)$ is the best (i.e. capacity-achieving) covariance matrix, $\lambda_i(\mathbf{R}_b) = [\mu - m / (\gamma \sigma_i^2(\mathbf{H}_0))]_+$ are the eigenvalues of \mathbf{R}_b (i.e. the optimum power allocation to the eigenmodes via water-filling), $i_+ = \{i : \lambda_i(\mathbf{R}) > 0\}$ is the index of active modes, and the constant μ is found from the total power constraint, $\sum_{i_+} \lambda_i(\mathbf{R}_b) = m$.

Proof: immediate by using (7), (8) and (12) in (10). Q.E.D.

Similarly to (12), the effect of channel uncertainty in this max-min problem is the SNR loss of $(1-\varepsilon)^2$. For example, a 3 dB loss occurs when $\varepsilon \approx 1 - 1/\sqrt{2} \approx 0.3$, i.e. 30% inaccuracy in the channel knowledge, and the condition for a negligible effect of uncertainty is $\varepsilon < 0.3$. As a side remark, we note that the capacity-achieving covariance has the following eigenvalue decomposition,

$$\mathbf{R}_b = \mathbf{V}_0 \Lambda_b \mathbf{V}_0{}^+ \quad (17)$$

where $\Lambda_b = \text{diag}(\lambda_1(\mathbf{R}_b) \dots \lambda_m(\mathbf{R}_b))$, the columns of \mathbf{V}_0 are the right singular vectors of \mathbf{H}_0 , which can be found from its SVD, $\mathbf{H}_0 = \mathbf{U}_0 \Sigma_0 \mathbf{V}_0{}^+$, i.e. the best Tx strategy is the transmission on the right singular vectors of the nominal channel, i.e. multiple beamforming, with the power distribution among the beams given by the water-filling algorithm applied to the nominal channel at the effective SNR $= (1-\varepsilon)^2 \gamma$. Under the best covariance in (17), the worst-case

channel perturbation in (15) takes the form,

$$\mathbf{E}_{\min} = -\mathbf{U}_0 \boldsymbol{\Sigma}_{\varepsilon} \mathbf{U}_0^+ \quad (18)$$

i.e. the eigenvectors of the worst perturbation are anti-parallel to the left singular vectors of the nominal channel and are of the maximum allowed amplitude. The intuition behind this result is that the jammer (uncertainty) tries to reduce the channel singular values as much as possible (since the capacity is a monotonically-increasing function of the singular values) and (16) is the best it can achieve (using (18)), under the best Tx strategy.

If the transmitter is not allowed to use non-uniform power allocation among the beams³, i.e. if $\lambda_1(\mathbf{R}) = \dots = \lambda_m(\mathbf{R}) = 1$, then $\mathbf{R}_b = \mathbf{V}_0 \mathbf{V}_0^+ = \mathbf{I}$ and

$$C_{\max \min} = I_{\min}(\mathbf{H}_0, I) \leq \max_{tr(\mathbf{R}) \leq m} I_{\min}(\mathbf{H}_0, \mathbf{R}).$$

Thus, the non-uniform power allocation is essential in achieving the capacity of the class of channels.

Theorem 2: For the class of channels in (2), (3), the capacity of the class equals to the worst-case channel capacity (the minimax and maxmin solutions are the same), i.e. there is a saddle point in $I(\mathbf{H}, \mathbf{R})$,

$$C_{\min \max} = C_{\max \min} \quad (19)$$

Proof: direct evaluation of $C_{\min \max}$, using the same approach as in Theorem 1. Q.E.D.

It follows that there exists a single code that achieves the capacity of the worst-case channel on any channel in the uncertainty class. Theorem 2 says in fact that in the matrix game between the transmitter (\mathbf{R}) and the jammer (\mathbf{E}) there is an optimum strategy for both players and each get penalized if it deviates from this strategy, provided that the other follows it. It also follows that the knowledge of \mathbf{E} does not help the transmitter to increase the capacity provided the jammer follows the best (i.e. capacity-minimizing) strategy, i.e. the optimization of \mathbf{R} for the nominal (\mathbf{H}_0) and true (\mathbf{H}) channels give the same capacity under the best jamming strategy (see (14)).

B. Additive uncertainty

The results of the previous section need only minor modifications to be adapted to the additive uncertainty model, so we give the proofs only when they are different.

Theorem 3: The capacity of the worst-case channel in the class in (4), (5) is

$$C_{\min \max} = \min_{\sigma_1(\Delta \mathbf{H}) \leq \varepsilon} \max_{tr(\mathbf{R}) \leq m} I(\mathbf{H}_0 + \Delta \mathbf{H}, \mathbf{R}) \\ = \sum_{i_+} \log \left(1 + \frac{\gamma}{m} (\sigma_i(\mathbf{H}_0) - \varepsilon)^2 \lambda_i(\mathbf{R}_b) \right) \quad (20)$$

where the summation is over $i_+ = \{i : \sigma_i(\mathbf{H}_0) > \varepsilon\}$, and $\lambda_i(\mathbf{R}_b) = [\mu - m / (\gamma(\sigma_i(\mathbf{H}_0) - \varepsilon)^2)]_+$ are the eigenvalues of the capacity-achieving covariance matrix \mathbf{R}_b .

Proof: applying the singular value inequalities in (13) to (8)

³ for example, due to the practical constrain of using the same fixed-power amplifiers

results in the lower bound on $C(\mathbf{H})$ equal to the right hand side of (20). Q.E.D.

We note that the worst-case channel perturbation $\Delta \mathbf{H}_w$ is given by

$$\Delta \mathbf{H}_w = -\mathbf{U}_0 \boldsymbol{\Sigma}_{\varepsilon_+} \mathbf{V}_0^+ \quad (21)$$

where $\boldsymbol{\Sigma}_{\varepsilon_+}$ is the diagonal matrix, with $\min\{\sigma_i(\mathbf{H}_0), \varepsilon\}$ on the main diagonal in position i , and the best covariance \mathbf{R}_b is the same as in (17). Thus, the worst channel perturbation is “anti-parallel” to the nominal channel, so that its singular values are reduced most.

Theorem 4: The capacity of the class of channels in (4), (5) is equal the capacity of the worst-case channel (the minimax and maxmin solutions are the same), i.e. there is a saddle point in $I(\mathbf{H}_0 + \Delta \mathbf{H}, \mathbf{R})$ under the additive channel uncertainty,

$$C_{\max \min} = C_{\min \max} \quad (22)$$

Proof: $C_{\max \min}$ can be evaluated using (8), (13) as follows,

$$C_{\max \min} = \max_{tr(\mathbf{R}) \leq m} \min_{\mathbf{H} \in S_H} I(\mathbf{H}, \mathbf{R}) \\ = \max_{tr(\mathbf{R}) \leq m} I(\mathbf{H}_w, \mathbf{R}) \\ = \max_{\sum_i \lambda_i(\mathbf{R}) \leq m} \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} \sigma_i^2(\mathbf{H}_w) \lambda_i(\mathbf{R}) \right) \\ \geq \max_{\sum_i \lambda_i(\mathbf{R}) \leq m} \sum_{i=1}^m \log \left(1 + \frac{\gamma}{m} (\sigma_i(\mathbf{H}_0) - \varepsilon)_+^2 \lambda_i(\mathbf{R}) \right) \\ = \sum_{i_+} \log \left(1 + \frac{\gamma}{m} (\sigma_i(\mathbf{H}_0) - \varepsilon)^2 \lambda_i(\mathbf{R}_b) \right) \\ = C_{\min \max} \quad (23)$$

where $\mathbf{H}_w = \mathbf{H}_w(\mathbf{H}_0, \mathbf{R}) = \arg \min I(\mathbf{H}, \mathbf{R})$ is the worst-case channel for given \mathbf{R} , and the inequality follows from (13). Combining (23) with (11) results in (22). Q.E.D.

Thus, there exists a single code that achieves the worst-case channel capacity for any channel in the uncertainty class. Neither the transmitter (\mathbf{R}) nor the jammer ($\Delta \mathbf{H}$) can deviate from the optimal strategy without incurring a loss. This result parallels one for the multiplicative uncertainty case. It can also be seen that the effect of uncertainty is to reduce each eigenmode gain from $\sigma_i(\mathbf{H}_0)^2$ to $(\sigma_i(\mathbf{H}_0) - \varepsilon)_+^2$, i.e. an SNR loss on each eigenmode individually. Unlike the case of multiplicative uncertainty model, this effect can not be expressed solely in terms of the aggregate SNR γ . A possible condition for the negligible effect of uncertainty is that $\varepsilon \ll \sigma_{\min}(\mathbf{H}_0)$, but it may be too conservative as small singular values may contribute little to the capacity. Note also that, in order to achieve the optimum, the transmitter needs to know only \mathbf{H}_0 (but not $\Delta \mathbf{H}$), since $\mathbf{R}_b = \mathbf{R}_b(\mathbf{H}_0)$ and the knowledge of $\Delta \mathbf{H}$ does not increase the capacity.

It should be noted that the results presented above reduce to the corresponding results in [12] (with the weight matrix $\mathbf{W} = \alpha \mathbf{I}$) when $\mathbf{H}_0 = \mathbf{h}_{r0} \mathbf{h}_{t0}^+$ is of rank one, where \mathbf{h}_{r0} and \mathbf{h}_{t0}^+ are the Rx and Tx array response vectors. In this case, $\mathbf{R}_b = m |\mathbf{h}_{t0}|^{-2} \mathbf{h}_{t0} \mathbf{h}_{t0}^+$, i.e. the best Tx strategy is the beamforming tailored to the nominal channel, and the worst-

case channel uncertainty is opposite of the nominal channel, $\Delta\mathbf{H}_w = -a\mathbf{H}_0$, where $a = \epsilon |\mathbf{h}_{r0}|^{-1} |\mathbf{h}_{t0}|^{-1}$. Thus, the spectral and Frobenius-normed channel uncertainties give the same result in this case.

Following the same approach as in [12], it can also be shown that the compound capacity above (and also the worst-case capacity, since both are equal) serves as a lower bound to the outage capacity C_{out} of a random non-ergodic channel with mean \mathbf{H}_0 and $\Delta\mathbf{H}$ representing the random part,

$$C_{out}(P_{out}) \geq C_{\max \min} \quad (24)$$

where the outage probability $P_{out} = \Pr\{\mathbf{H} \notin S_H\}$, i.e. the probability that channel realization is not in the uncertainty set. Furthermore, if the outage capacity is defined for the same Tx covariance as the compound capacity, then the equality is achieved in (24). Thus, the compound capacity can also serve as the outage capacity at appropriately defined outage probability.

V. CONCLUSION

The capacity of a class of uncertain MIMO channels (compound capacity) subject to the spectral norm constraint has been considered for both multiplicative and additive uncertainty models. The compound capacity, achievable by a single code for all channels in the class, has been shown to be equal to the worst-case channel capacity, via explicit closed-form evaluation of the two. Thus, the saddle-point property holds and, in terms of the game theory, neither player can deviate from the optimum strategy without incurring loss. The best transmission strategy and the worst-case channel uncertainties are given in a simple form: the former is the transmission on the eigenmodes of the nominal channel (i.e. multiple beamforming + water filling to distribute the Tx power among the beams), and the latter is opposite of the nominal channel. The effect of uncertainty is equivalent to an SNR loss commensurable with the size of the uncertainty set. In the case of rank-one channel, the best Tx strategy is beamforming on the nominal channel, which is in agreement with [12]. The compound channel capacity serves as a lower bound on the outage capacity.

VI. ACKNOWLEDGEMENT

The authors would like to thank F. Kschischang, S. Denis and

N. Ahmed for many useful discussions.

VII. REFERENCES

- [1] I. E. Telatar, "Capacity of Multi-Antenna Gaussian Channels", AT&T Bell Labs, Internal Tech. Memo, June 1995, (European Trans. Telecom., v.10, no. 6, Dec. 1999).
- [2] G. J. Foschini and M. J. Gans, "On Limits of Wireless Communications in a Fading Environment when using Multiple Antennas", *Wireless Personal Commun.*, vol.6, no. 3, pp. 311-335, March 1998.
- [3] H. Bolcskei et al (Eds.), *Space-Time Wireless Systems: From Array Processing to MIMO Communications*, Cambridge University Press, Cambridge, 2006.
- [4] D. Blackwell et al, *The Capacity of a Class of Channels*, *Ann. Math. Statist.*, v. 30, pp. 1229-1241, 1959.
- [5] W.L. Root, P.P. Varaiya, *Capacity of Classes of Gaussian Channels*, *SIAM J. Appl. Math.*, v. 16, N. 6, pp. 1350-1393, Nov. 1968.
- [6] R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
- [7] T.L. Marzetta, B.M. Hochwald, *Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading*, *IEEE Transactions on Information Theory*, v. 45, N. 1, pp. 139-157, Jan 1999.
- [8] A. Lapidoth, P. Narayan, *Reliable Communications Under Channel Uncertainty*, *IEEE Trans. Information Theory*, v. 44, N. 6, pp. 2148-2177, Oct. 1998.
- [9] M. Medard, *The Effect upon Channel Capacity in Wireless Communications of Perfect and Imperfect Knowledge of the Channel*, *IEEE Trans. Information Theory*, v. 46, N. 3, pp. 933-946, May 2000.
- [10] T. Weber et al, *Imperfect Channel-State Information in MIMO Transmission*, *IEEE Trans. Communications*, v. 54, N. 3, pp. 543-552, Mar. 2006.
- [11] A. Pascual-Iserte et al, *A Robust Maxmin Approach for MIMO Communications With Imperfect Channel State Information Based on Convex Optimization*, *IEEE Trans. Signal Processing*, v. 54, N. 1, pp. 346-360, Jan. 2006.
- [12] A. Wiesel et al, *Optimization of the MIMO Compound Capacity*, *IEEE Trans. Wireless Communications*, v. 6, N. 3, pp. 1094-1101, Mar. 2007.
- [13] C.D. Charalambous, S.Z. Denic and S. Djouadi, "Robust Capacity of White Noise Channels with Uncertainty," in *Proceedings of the 43rd IEEE Conference on Decision and Control*, pp.4880-4884, Bahamas, December 14-17, 2004.