Capacity of Compound MIMO Gaussian Channels with Additive Uncertainty

Yin Sun, C. Emre Koksal, and Ness B. Shroff
Dept. of ECE, The Ohio State University, Columbus, OH

***Abstract*—This paper considers reliable communications over a multiple-input multiple-output (MIMO) Gaussian channel, where the channel matrix is within a bounded channel uncertainty region around a nominal channel matrix, i.e., an instance of the compound MIMO Gaussian channel. We study the optimal transmit covariance design to achieve the capacity of compound MIMO Gaussian channels, where the channel uncertainty region is characterized by the spectral norm. This design problem is a challenging non-convex optimization problem. However, in this paper, we reveal that this design problem has a *hidden convexity* property, and hence it can be simplified as a convex optimization problem. Towards this goal, we first prove that the optimal transmit design is to diagonalize the nominal channel, and then show that the duality gap between the capacity of the compound MIMO Gaussian channel and the minimal channel capacity is zero, which proves the conjecture of Loyka and Charalambous (*IEEE Trans. Inf. Theory*, vol. 58, no. 4, pp. 2048-2063, 2012). The key tools for showing these results are a novel matrix determinant inequality and some unitarily invariant properties.**

I. Introduction

Multiple-input multiple-output (MIMO) techniques have been prevalent to improve the spectral efficiencies of wireless communications. The performance of MIMO communications relies on access to the channel state information (CSI). When the CSI is perfectly known at the transmitter, the optimal power allocation is to diagonalize the channel [1]. However, in practice, the transmitter often has some channel uncertainty, which can result in a significant rate loss, if not taken into consideration in the transmit covariance design.

There has been two categories of research towards reliable communications over MIMO Gaussian channels with channel uncertainty. The first category focuses on stochastic models of channel uncertainty, where the transmitter has access to only the statistics of the channel state, but not to its realization. When the channel states change quickly over time, the achievable rate of the channel is described by the ergodic capacity, e.g., [1]–[4]. On the other hand, when the channel states vary slowly, the achievable rate is characterized by the outage capacity, which is the maximum data rate achievable at any given state with probability no smaller than a specified value, e.g., [1], [3]–[7].

The second category of studies were centered on deterministic models of channel uncertainty, where the CSI is a deterministic variable within a known set, but its actual value is unknown to the transmitter. Such a model is called a *compound channel* in information theory [8]. From a practical viewpoint, it is the maximum data rate that can be reliably transmitted over *any* channel from the given set. Characterizing the capacity of the compound channel is considered to be an important problem, and has received considerable attention.

In closed-loop MIMO systems, the transmitter is able to obtain an inaccurate CSI, where the channel error may be caused by estimation, interpolation, mobility, and/or feedback. In this case, the channel is typically modelled as the sum of a known nominal channel and an unknown channel uncertainty. This additive channel uncertainty model has been widely utilized both in information theoretical studies, e.g., [9]–[11], and in the robust transceiver designs in signal processing literature, e.g., [12]–[15]. In [16], the capacity of the compound Rician MIMO Gaussian channel with additive channel uncertainty was studied, where the analysis was restricted to a rank-one nominal channel. Arbitrary rank nominal channel was considered in [17], where the channel uncertainty is limited to the singular value of the nominal channel with no uncertainty on the singular vectors. The capacity of the compound MIMO channel with a multiplicative channel uncertainty model was obtained in [11], where the region of channel uncertainty is described by spectral norm. In addition, the capacity of the compound MIMO Gaussian channel with additive channel uncertainty was derived in [11] for some special cases, such as high signal-to-noise ratio (SNR) limit, low SNR limit, and rank-two nominal channel.

In this paper, we design the optimal transmit covariance to achieve the capacity of the compound MIMO Gaussian channel with additive channel uncertainty. We consider the case that the channel uncertainty is in a bounded region around the nominal channel matrix, which is characterized by the spectral norm. This design problem is a challenging nonconvex optimization problem. However, we reveal that the transmit covariance design problem possesses a *hidden convexity* property, and hence it can be simplified as a convex optimization problem. We first prove that the optimal transmit covariance design is to diagonalize the nominal channel. Then, we show that the duality gap between the capacity of the compound MIMO Gaussian channel and the minimal channel capacity is zero, which proves the conjecture of Loyka and Charalambous [11]. The key tools for proving these results are a novel matrix determinant inequality (Lemma 1) and some unitarily invariant properties.

II. System Model

A. Notation

The following notations are used. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and standard lower case letters denote scalars. Let $\mathbb{C}^{m\times n}$ denote the set of $m\times n$ complex-valued matrices, and $\mathbb{C}^n$ denote the set of $n\times n$ square complex-valued matrices. The symbol $\mathbb{S}^n$ represents the set of $n\times n$ Hermitian matrices, and $\mathbb{S}^n_+$ represents the set of $n\times n$ Hermitian positive semidefinite matrices. The operator $\text{diag}(x_1, x_2, \cdots, x_n)$ denotes a diagonal matrix with diagonal entries given by $x_1, x_2, \cdots, x_n$. The matrix $\mathbf{I}_n$ denotes the $n\times n$ identity matrix. By $\boldsymbol{x}\geq \mathbf{0}$, we mean that $x_i\geq 0$ for all i. The operators $(\cdot)^H$, $\text{Tr}(\cdot)$ and $\det(\cdot)$ on matrices denote the Hermitian, trace and determinant operations, respectively. Let $\sigma_i(\mathbf{A})$ and $\lambda_i(\mathbf{A})$ represent the singular value and eigenvalue of $\mathbf{A}$, respectively. The vector $\boldsymbol{\sigma}(\mathbf{A})\triangleq(\sigma_1(\mathbf{A}),\cdots,\sigma_{\min\{m,n\}}(\mathbf{A}))$ contains the singular values of $\mathbf{A}\in\mathbb{C}^{m\times n}$. Let $\boldsymbol{\lambda}(\mathbf{Q})\triangleq(\lambda_1(\mathbf{Q}),\cdots,\lambda_n(\mathbf{Q}))$ denote a vector containing the eigenvalues of $\mathbf{Q}\in\mathbb{S}^n$. The singular values and eigenvalues are listed in descending order. We use $|||\cdot|||$ and $\|\cdot\|$ to denote matrix norm and vector norm, respectively.

B. Channel Model

Consider the complex-valued Gaussian vector channel:

$$\boldsymbol{y}=\mathbf{H}\boldsymbol{x}+\mathbf{n}, \tag{1}$$

where $\boldsymbol{y}$ is a length r received vector, $\mathbf{H}$ is a $r\times t$ channel matrix, $\boldsymbol{x}$ is a length t transmitted vector with zero mean and covariance $E\{\boldsymbol{x}\boldsymbol{x}^H\}=\mathbf{Q}$, and $\mathbf{n}$ is a complex Gaussian noise vector with zero-mean and covariance $E\{\mathbf{n}\mathbf{n}^H\}=\mathbf{I}_r$.

The MIMO channel $\mathbf{H}$ is an unknown deterministic matrix satisfying

$$\mathbf{H}\in\mathcal{H}, \tag{2}$$

where $\mathcal{H}$ is the channel uncertainty region defined by

$$\mathcal{H}\triangleq\{\mathbf{H}: |||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon\}, \tag{3}$$

$\mathbf{H}_0$ is the nominal channel, and $|||\cdot|||_2$ is the spectral norm defined by

$$|||\,\mathbf{A}|||_2\triangleq\max_{\|\boldsymbol{x}\|_2\leq 1}\|\mathbf{A}\boldsymbol{x}\|_2=\max_i\{\sigma_i(\mathbf{A})\}=\|\boldsymbol{\sigma}(\mathbf{A})\|_\infty. \tag{4}$$

The spectral norm is a *unitarily invariant matrix norm.* A unitarily invariant matrix norm satisfies [18, Section 7.4.16]

$$|||\,\mathbf{U}\mathbf{A}\mathbf{V}|||=|||\mathbf{A}||| \tag{5}$$

for all $\mathbf{A}\in\mathbb{C}^{m\times n}$ and for all unitary matrices $\mathbf{U}\in\mathbb{C}^m$ and $\mathbf{V}\in\mathbb{C}^n$. Therefore, the channel uncertainty $\boldsymbol{\Delta}=\mathbf{H}-\mathbf{H}_0$ is within an isotopical set. Note that the channel uncertainty region (3) provides a conservative performance lower bound for the regions defined by any other unitarily invariant matrix norm, because

$$|||\mathbf{A}|||_2\geq|||\mathbf{A}|||$$

holds for all matrix $\mathbf{A}$ and all unitarily invariant matrix norm $|||\cdot|||$ [18, Corollary 5.6.35]. More discussions about the relationship among different matrix norms are provided in [19].

C. Power Constraint

We consider a general transmit power constraint

$$\mathbf{Q}\in\mathcal{Q}, \tag{6}$$

where $\mathcal{Q}\subset\mathbb{S}^t_+$ is a nonempty compact convex set satisfying

$$\mathbf{U}\mathbf{Q}\mathbf{U}^H\in\mathcal{Q}, \tag{7}$$

$$\mathbf{D}(\mathbf{Q})\in\mathcal{Q}, \tag{8}$$

for all $\mathbf{Q}\in\mathcal{Q}$ and all unitary matrix $\mathbf{U}\in\mathbb{C}^t$, where $\mathbf{D}(\mathbf{Q})$ is the diagonal matrix with the same diagonal elements with $\mathbf{Q}$. We say a set $\mathcal{Q}$ is *unitarily invariant* if it satisfies (7) and (8). One can show that each unitarily invariant $\mathcal{Q}$ can be equivalently expressed as

$$\mathcal{Q}=\{\mathbf{Q}\in\mathbb{S}^t_+:\boldsymbol{\lambda}(\mathbf{Q})\in B_{\mathcal{Q}},\boldsymbol{\lambda}(\mathbf{Q})\geq\mathbf{0}\}, \tag{9}$$

where $B_{\mathcal{Q}}$ is a nonempty compact convex set. Two familiar examples of unitarily invariant power constraints are the sum power constraint [1]

$$\begin{aligned}\mathcal{Q}_1&=\{\mathbf{Q}\in\mathbb{S}^t_+:\text{Tr}(\mathbf{Q})\leq t\}, \qquad (10)\\ &=\{\mathbf{Q}\in\mathbb{S}^t_+:\sum_{i=1}^t\lambda_i(\mathbf{Q})\leq t,\boldsymbol{\lambda}(\mathbf{Q})\geq\mathbf{0}\},\end{aligned}$$

and the maximum power constraint [14]

$$\mathcal{Q}_2=\{\mathbf{Q}\in\mathbb{S}^t_+:\max_i\{\lambda_i(\mathbf{Q})\}\leq P_m,\boldsymbol{\lambda}(\mathbf{Q})\geq\mathbf{0}\}.$$

III. Optimal Transmit Covariance Design

A. Main Result

The capacity of the compound MIMO Gaussian channel (1)-(3) and (6) is [20, Theorem 7.1]

$$C_{\max\min}\triangleq\max_{\mathbf{Q}\in\mathcal{Q}}\min_{|||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon}I(\mathbf{Q},\mathbf{H}), \tag{11}$$

where $I(\mathbf{Q},\mathbf{H})=I(\boldsymbol{x};\boldsymbol{y})$ is the mutual information of the channel (1), i.e., [1]

$$I(\mathbf{Q},\mathbf{H})=\log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right),$$

and γ is the per-antenna SNR. Efficient solution of the max-min problem (11) has been open for a long time (except for some special cases [11], [14], [16]), because $I(\mathbf{Q},\mathbf{H})$ is nonconvex with respect to $\mathbf{H}$.

Suppose that the singular value decomposition (SVD) of the nominal channel $\mathbf{H}_0$ is given by

$$\mathbf{H}_0=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}_0}\mathbf{V}_0^H, \tag{12}$$

where $\mathbf{U}_0\in\mathbb{C}^r$ and $\mathbf{V}_0\in\mathbb{C}^t$ are unitary matrices. The first key result of this paper is stated as follows:

Theorem 1. *If $\mathcal{Q}$ and $\mathcal{H}$ are nonempty sets, $\mathcal{H}$ is defined in* (3)*, and $\mathcal{Q}$ satisfies the unitarily invariant properties* (7) *and* (8)*, then*

$$\mathbf{Q}^\star=\mathbf{V}_0\boldsymbol{\Lambda}^\star_{\mathbf{Q}}\mathbf{V}_0^H,\ \mathbf{H}^\star=\mathbf{U}_0\boldsymbol{\Sigma}^\star_{\mathbf{H}}\mathbf{V}_0^H, \tag{13}$$

is a solution to Problem (11)*, where $\mathbf{U}_0$ and $\mathbf{V}_0$ are defined in* (12)*, the diagonal matrices $\boldsymbol{\Lambda}^\star_{\mathbf{Q}}$ and $\boldsymbol{\Sigma}^\star_{\mathbf{H}}$ are determined* *by $\mathbf{\Sigma}_{\mathbf{H}}^{\star} = diag(\boldsymbol{\sigma}^{\star})$ and $\mathbf{\Lambda}_{\mathbf{Q}}^{\star} = diag(\boldsymbol{\lambda}^{\star})$, such that $(\boldsymbol{\sigma}^{\star}, \boldsymbol{\lambda}^{\star})$ solves the problem*

$$C_{\max\min} = \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}} \\ \boldsymbol{\lambda}\geq \mathbf{0}}} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|_\infty\leq\varepsilon \\ \boldsymbol{\sigma}\geq\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i), \quad (14)$$

with the convex set $B_{\mathcal{Q}}$ defined in (9).

Proof: The proof of Theorem 1 relies on the unitarily invariant properties (4), (5), (7), and (8), and a novel matrix determinant inequality presented in Lemma 1 given below. The proof details are provided in Appendix A. ■

The following lemma is an important technical contribution of this paper, which plays a key role in proving Theorem 1.

Lemma 1 (Matrix Determinant Inequality)**.** *If $\mathbf{\Sigma}$ and $\mathbf{\Lambda}$ are diagonal matrices with nonnegative diagonal entries, then one solution to*

$$\min_{|||\mathbf{\Delta}|||_2\leq\varepsilon} \det\left[\mathbf{I}+(\mathbf{\Sigma}+\mathbf{\Delta})\mathbf{\Lambda}(\mathbf{\Sigma}+\mathbf{\Delta})^H\right] \quad (15)$$

is a diagonal matrix.

The proof of Lemma 1 is provided in [19].

Theorem 1 implies that the optimal transmit covariance of the MIMO Gaussian channel with worst case channel uncertainty is to diagonalize the nominal channel $\mathbf{H}_0$, if the conditions in Theorem 1 are satisfied. Such a solution structure was previously known only for some special cases, such as high SNR limit ($\gamma \gg 1$), low SNR limit ($\gamma \ll 1$), low rank nominal channels (rank$(\mathbf{H}_0) \leq 2$) [11], [14], [16], while Theorem 1 holds for general nominal channels and all SNR values. By Theorem 1, the problem (11) reduces to (14) with much fewer variables.

B. The Dual Problem

Now we consider the duality to the max-min problem (11): the minimal capacity of the MIMO Gaussian channels, given by the following min-max problem

$$C_{\min\max} \triangleq \min_{\mathbf{H}\in\mathcal{H}} \max_{\mathbf{Q}\in\mathcal{Q}} I(\mathbf{Q},\mathbf{H}). \quad (16)$$

It is important to distinguish the capacity of the compound channel $C_{\max\min}$ and the minimal channel capacity $C_{\min\max}$: $C_{\max\min}$ can be achieved for any channel $\mathbf{H}$ within $\mathcal{H}$, by using the same transmit covariance matrix $\mathbf{Q}$.[1] $C_{\min\max}$ is the minimal capacity of the channels with $\mathbf{H}\in\mathcal{H}$, evaluating which requires knowledge of $\mathbf{H}$ at the transmitter to obtain $\mathbf{Q}$.[2] We study the min-max problem (16) to gain more insight into the max-min problem (11). We consider a more general channel uncertainty region

$$\mathcal{H} \triangleq \{\mathbf{H} : |||\mathbf{H}-\mathbf{H}_0||| \leq \varepsilon\}, \quad (17)$$

where $|||\cdot|||$ is a unitarily invariant matrix norm satisfying (5). For any unitarily invariant matrix norm $|||\cdot|||$, there is a vector norm $\|\cdot\|$ such that

$$|||\,\mathbf{A}||| = \|\boldsymbol{\sigma}(\mathbf{A})\| \quad (18)$$

holds for all $\mathbf{A}\in\mathbb{C}^{m\times n}$ [21, Theorem 3.5.18]. For the special case of spectral norm, the associated vector norm in (18) is $\|\cdot\|_\infty$, as given by (4). We have the following result:

Theorem 2. *If $\mathcal{Q}$ and $\mathcal{H}$ are nonempty sets, $\mathcal{H}$ is defined in* (17)*, and $\mathcal{Q}$ satisfies the unitarily invariant properties* (7) *and* (8)*, then*

$$\mathbf{Q}' = \mathbf{V}_0\mathbf{\Lambda}_{\mathbf{Q}}'\mathbf{V}_0^H,\ \mathbf{H}' = \mathbf{U}_0\mathbf{\Sigma}_{\mathbf{H}}'\mathbf{V}_0^H, \quad (19)$$

is a solution to Problem (16)*, where $\mathbf{U}_0$ and $\mathbf{V}_0$ are defined in* (12)*, the diagonal matrices $\mathbf{\Lambda}_{\mathbf{Q}}'$ and $\mathbf{\Sigma}_{\mathbf{H}}'$ are determined by $\mathbf{\Sigma}_{\mathbf{H}}' = diag(\boldsymbol{\sigma}')$ and $\mathbf{\Lambda}_{\mathbf{Q}}' = diag(\boldsymbol{\lambda}')$ such that $(\boldsymbol{\sigma}', \boldsymbol{\lambda}')$ solves the problem*

$$C_{\min\max} = \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|\leq\varepsilon \\ \boldsymbol{\sigma}\geq\mathbf{0}}} \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}} \\ \boldsymbol{\lambda}\geq \mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i), \quad (20)$$

with the vector norm $\|\cdot\|$ and the convex set $B_{\mathcal{Q}}$ defined in (18) *and* (9)*, respectively.*

Proof: The proof of Theorem 2 relies on the unitarily invariant properties (18), (5), (7), and (8), but not the matrix determinant inequality in Lemma 1 for the spectral norm case. Therefore, Theorem 2 holds for any unitarily invariant matrix norm. The proof details are provided in Appendix B. ■

Note that a special case of Theorem 2 was obtained in Theorem 3 of [11], where $|||\cdot|||$ is limited to the spectral norm $|||\cdot|||_2$ and $\mathcal{Q}$ is the sum power constraint $\mathcal{Q}_1$.

C. Duality Gap is Zero

It is interesting to see that the max-min problem (11) and the min-max problem (16) have similar solution structures, as given in (13) and (19), and the difference is only in the solutions to (14) and (20). Next, we study whether (14) and (20) have a common solution for the spectral norm case.

It is known that the following weak duality relation is always true: [20]

$$C_{\max\min} \leq C_{\min\max}. \quad (21)$$

Moreover, equality holds in (21), i.e.,

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

if and only if (14) and (20) have a common solution [22, Corollary 9.16]. It was conjectured in [11] that (22) holds for the case that $|||\cdot||| = |||\cdot|||_2$ and the power constraint is $\mathcal{Q} = \mathcal{Q}_1$. Here, using Theorem 1 and 2, we can simply prove this conjecture:

Theorem 3. *If the conditions of Theorem 1 are satisfied, then:*

1) *The strong duality relation* (22) *holds.*
2) *Problems* (14) *and* (20) *have a common solution $(\boldsymbol{\sigma}^*, \boldsymbol{\lambda}^*)$, where $\boldsymbol{\sigma}^*$ is given by*

$$\sigma_i^* = \max\{\sigma_{0,i}-\varepsilon, 0\}, \quad (23)$$

and $\boldsymbol{\lambda}^*$ is determined by the convex optimization problem

$$C_{\max\min} = \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}} \\ \boldsymbol{\lambda}\geq \mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\max\{\sigma_{0,i}-\varepsilon,0\}^2\lambda_i). \quad (24)$$

[1]The outer optimization of $\mathbf{Q}$ in (11) is done without knowledge of $\mathbf{H}$.
[2]The inner optimization of $\mathbf{Q}$ in (16) is done with knowledge of $\mathbf{H}$.

Proof: 1) Problem (14) can be expressed as

$$\max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\geq \mathbf{0}}} \min_{\boldsymbol{\sigma}\geq\mathbf{0}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i) \tag{25}$$
$$\text{s.t.} \quad \max\{\sigma_{0,i}-\varepsilon,0\}\leq\sigma_i\leq\sigma_{0,i}+\varepsilon,\ \forall\ i.$$

By introducing $x_i \triangleq \log(\sigma_i)$, this problem can be reformulated as the following convex optimization problem:

$$\max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\geq \mathbf{0}}} \min_{\boldsymbol{x}} \sum_{i=1}^{\min\{t,r\}} \log[1+\gamma e^{2x_i}\lambda_i]$$
$$\text{s.t.} \quad \log(\max\{\sigma_{0,i}-\varepsilon,0\})\leq x_i\leq\log(\sigma_{0,i}+\varepsilon),\ \forall\ i,$$

where the objective function is concave in $\boldsymbol{\lambda}$ and convex in $\boldsymbol{x}$ [23]. Similarly, (20) can be also reformulated as a convex optimization problem. Then, (22) follows from von Neumann's minimax theorem [22, Theorem 9.D].

2) For any $\boldsymbol{\lambda}$, the inner minimization problem of (25) can be separated into several subproblems, and the solution is given by (23). Thus, Problem (25) reduces to (24). ∎

We note that the conjecture of [11] is a special case of Theorem 3 where $\mathcal{Q}$ is restricted to be the sum power constraint $\mathcal{Q}_1$. By Theorem 1 and 3, we have shown that the covariance design problem (11) is a convex optimization problem in nature, if the channel uncertainty region $\mathcal{H}$ is characterized by the spectral norm.

IV. Conclusion

In this paper, we have investigated the capacity of a compound MIMO channel with an additive uncertainty of bounded spectral norm, and derived the optimal transmit covariance matrix in close-form. When the channel uncertainty region is characterized by the spectral norm, we have revealed a hidden convexity property in this problem. We have proved that the optimal transmit covariance design is to diagonalize the nominal channel matrix and there is zero duality gap between the capacity of the compound MIMO Gaussian channel and the minimal channel capacity.

Appendix A
Proof of Theorem 1

First, we construct an upper bound of $C_{\max\min}$ by imposing one extra constraint in the inner minimization problem:

$$\begin{aligned} &C_{\max\min}\\ &\overset{(a)}{\leq} \max_{\mathbf{Q}\in\mathcal{Q}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon\\ \mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\\ &\overset{(b)}{=} \max_{\mathbf{Q}\in\mathcal{Q}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon\\ \mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H\mathbf{Q}\mathbf{V}_0\boldsymbol{\Sigma}_{\mathbf{H}}\right), \end{aligned}$$

where $\mathbf{U}_0$ and $\mathbf{V}_0$ are defined in (12), step (a) is due to the additional constraint in the inner minimization problem, and step (b) is due to $\mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H$ and $\det(\mathbf{I}+\mathbf{AB})=\det(\mathbf{I}+\mathbf{BA})$. Let us define $\tilde{\mathbf{Q}}\triangleq\mathbf{V}_0^H\mathbf{Q}\mathbf{V}_0$ and use $\mathbf{D}(\tilde{\mathbf{Q}})$ to denote the diagonal matrix that has the same diagonal elements with $\tilde{\mathbf{Q}}$, then we attain

$$\begin{aligned} &C_{\max\min}\\ &\leq \max_{\mathbf{Q}\in\mathcal{Q}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon\\ \mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\tilde{\mathbf{Q}}\boldsymbol{\Sigma}_{\mathbf{H}}\right)\\ &\overset{(a)}{=} \max_{\tilde{\mathbf{Q}}\in\mathcal{Q}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\leq\varepsilon\\ \mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\tilde{\mathbf{Q}}\boldsymbol{\Sigma}_{\mathbf{H}}\right)\\ &\overset{(b)}{=} \max_{\tilde{\mathbf{Q}}\in\mathcal{Q}} \min_{|||\boldsymbol{\Sigma}_{\mathbf{H}}-\boldsymbol{\Sigma}_{\mathbf{H}_0}|||_2\leq\varepsilon} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\tilde{\mathbf{Q}}\boldsymbol{\Sigma}_{\mathbf{H}}\right)\\ &\overset{(c)}{\leq} \max_{\tilde{\mathbf{Q}}\in\mathcal{Q}} \min_{|||\boldsymbol{\Sigma}_{\mathbf{H}}-\boldsymbol{\Sigma}_{\mathbf{H}_0}|||_2\leq\varepsilon} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{D}(\tilde{\mathbf{Q}})\boldsymbol{\Sigma}_{\mathbf{H}}\right)\\ &\overset{(d)}{\leq} \max_{\mathbf{D}(\tilde{\mathbf{Q}})\in\mathcal{Q}} \min_{|||\boldsymbol{\Sigma}_{\mathbf{H}}-\boldsymbol{\Sigma}_{\mathbf{H}_0}|||_2\leq\varepsilon} \log\det\left(\mathbf{I}_r+\gamma\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{D}(\tilde{\mathbf{Q}})\boldsymbol{\Sigma}_{\mathbf{H}}\right)\\ &\overset{(e)}{=} \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\geq\mathbf{0}}} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|_\infty\leq\varepsilon\\ \boldsymbol{\sigma}\geq\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i), \end{aligned} \tag{26}$$

where step (a) is due to (7), step (b) is due to $|||\boldsymbol{\Sigma}_{\mathbf{H}}-\boldsymbol{\Sigma}_{\mathbf{H}_0}|||_2=|||\mathbf{H}-\mathbf{H}_0|||_2$ that is derived from $\mathbf{H}_0=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}_0}\mathbf{V}_0^H$, $\mathbf{H}=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H$ and (5), step (c) is due to the Hadamard inequality $\det(\mathbf{A})\leq\prod_i\mathbf{A}_{ii}$, step (d) is due to that the feasible region $\mathbf{D}(\tilde{\mathbf{Q}})\in\mathcal{Q}$ is larger than the region $\tilde{\mathbf{Q}}\in\mathcal{Q}$ according to (8), and step (e) is due to (4) and (9) with λ_i representing the diagonal entries of $\mathbf{D}(\tilde{\mathbf{Q}})$.

Next, we build a lower bound of $C_{\max\min}$ by considering one extra constraint in the outer maximization problem:

$$\begin{aligned} &C_{\max\min}\\ &\overset{(a)}{\geq} \max_{\substack{\mathbf{Q}\in\mathcal{Q}\\ \mathbf{Q}=\mathbf{V}_0\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{V}_0^H}} \min_{\mathbf{H}\in\mathcal{H}} I(\mathbf{Q},\mathbf{H})\\ &= \max_{\substack{\mathbf{Q}\in\mathcal{Q}\\ \mathbf{Q}=\mathbf{V}_0\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{V}_0^H}} \min_{|||\boldsymbol{\Delta}|||_2\leq\varepsilon} \log\det\left[\mathbf{I}_r+\gamma(\mathbf{H}_0+\boldsymbol{\Delta})\mathbf{Q}(\mathbf{H}_0+\boldsymbol{\Delta})^H\right]\\ &\overset{(b)}{=} \max_{\boldsymbol{\Lambda}_{\mathbf{Q}}\in\mathcal{Q}} \min_{|||\tilde{\boldsymbol{\Delta}}|||_2\leq\varepsilon} \log\det\left[\mathbf{I}_r+\gamma(\boldsymbol{\Sigma}_{\mathbf{H}_0}+\tilde{\boldsymbol{\Delta}})\boldsymbol{\Lambda}_{\mathbf{Q}}(\boldsymbol{\Sigma}_{\mathbf{H}_0}+\tilde{\boldsymbol{\Delta}})^H\right], \end{aligned} \tag{27}$$

where $\tilde{\boldsymbol{\Delta}}\triangleq\mathbf{U}_0^H\boldsymbol{\Delta}\mathbf{V}_0$, step (a) is due the additional constraint in the outer maximization, and step (b) is due to $\mathbf{H}_0=\mathbf{U}_0\boldsymbol{\Sigma}_{\mathbf{H}_0}\mathbf{V}_0^H$, $\mathbf{Q}=\mathbf{V}_0\boldsymbol{\Lambda}_{\mathbf{Q}}\mathbf{V}_0^H$, the definition $\tilde{\boldsymbol{\Delta}}\triangleq\mathbf{U}_0^H\boldsymbol{\Delta}\mathbf{V}_0$, and the unitarily invariant properties (5) and (7).

According to Lemma 1, the optimal $\tilde{\boldsymbol{\Delta}}$ is a diagonal matrix. Hence, $\boldsymbol{\Sigma}'_{\mathbf{H}}=\boldsymbol{\Sigma}_{\mathbf{H}_0}+\tilde{\boldsymbol{\Delta}}$ in (27) is also a diagonal matrix. Substituting this into (27), we have

$$\begin{aligned} C_{\max\min} &\geq \max_{\boldsymbol{\Lambda}_{\mathbf{Q}}\in\mathcal{Q}} \min_{|||\boldsymbol{\Sigma}'_{\mathbf{H}}-\boldsymbol{\Sigma}_{\mathbf{H}_0}|||_2\leq\varepsilon} \log\det\left[\mathbf{I}_r+\gamma\boldsymbol{\Sigma}'_{\mathbf{H}}\boldsymbol{\Lambda}_{\mathbf{Q}}\boldsymbol{\Sigma}'_{\mathbf{H}}\right]\\ &= \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\geq\mathbf{0}}} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|_\infty\leq\varepsilon\\ \boldsymbol{\sigma}\geq\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i), \end{aligned} \tag{28}$$

where the last step is due to (4) and (9) with σ_i representing the diagonal entries of $\boldsymbol{\Sigma}'_{\mathbf{H}}$. Using (26) and (28), the optimal objective value of (11) is given by (14).

Finally, we show (13) is an optimal solution to (11). For this, we substitute the solution (13) into (11), i.e.,

$$\begin{aligned}
&\max_{\substack{\mathbf{Q}\in\mathcal{Q}\\ \mathbf{Q}=\mathbf{V}_0\mathbf{\Lambda}_{\mathbf{Q}}^\star\mathbf{V}_0^H}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\le\varepsilon\\ \mathbf{H}=\mathbf{U}_0\mathbf{\Sigma}_{\mathbf{H}}^\star\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\\
&= \max_{\substack{\mathbf{Q}\in\mathcal{Q}\\ \mathbf{Q}=\mathbf{V}_0\mathbf{\Lambda}_{\mathbf{Q}}^\star\mathbf{V}_0^H}} \min_{\substack{|||\mathbf{H}-\mathbf{H}_0|||_2\le\varepsilon\\ \mathbf{H}=\mathbf{U}_0\mathbf{\Sigma}_{\mathbf{H}}^\star\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{\Sigma}_{\mathbf{H}}^\star\mathbf{\Lambda}_{\mathbf{Q}}^\star\mathbf{\Sigma}_{\mathbf{H}}^\star\right)\\
&\overset{(a)}{=} \max_{\mathbf{\Lambda}_{\mathbf{Q}}^\star\in\mathcal{Q}} \min_{|||\mathbf{\Sigma}_{\mathbf{H}}^\star-\mathbf{\Sigma}_{\mathbf{H}_0}|||_2\le\varepsilon} \log\det\left(\mathbf{I}_r+\gamma\mathbf{\Sigma}_{\mathbf{H}}^\star\mathbf{\Lambda}_{\mathbf{Q}}^\star\mathbf{\Sigma}_{\mathbf{H}}^\star\right)\\
&\overset{(b)}{=} \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\ge\mathbf{0}}} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|_\infty\le\varepsilon\\ \boldsymbol{\sigma}\ge\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i)\\
&= C_{\max\min},
\end{aligned} \tag{29}$$

where step (a) is due to (5) and (7), step (b) is due to (9) and (18). By this, the theorem is proved. ■

Appendix B
Proof of Theorem 2

Consider the following upper bound of $C_{\min\max}$:

$$\begin{aligned}
C_{\min\max} &\le \min_{\substack{\mathbf{H}\in\mathcal{H}\\ \mathbf{H}=\mathbf{U}_0\mathbf{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \max_{\mathbf{Q}\in\mathcal{Q}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\\
&\overset{(a)}{=} \min_{\substack{\mathbf{H}\in\mathcal{H}\\ \mathbf{H}=\mathbf{U}_0\mathbf{\Sigma}_{\mathbf{H}}\mathbf{V}_0^H}} \max_{\substack{\mathbf{Q}\in\mathcal{Q}\\ \mathbf{Q}=\mathbf{V}_0\mathbf{\Lambda}_{\mathbf{Q}}\mathbf{V}_0^H}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\\
&\overset{(b)}{=} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|\le\varepsilon\\ \boldsymbol{\sigma}\ge\mathbf{0}}} \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\ge\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i),
\end{aligned} \tag{30}$$

where step (a) is due to that the optimal power allocation result is of the form $\mathbf{Q}=\mathbf{V}_0\mathbf{\Lambda}_{\mathbf{Q}}\mathbf{V}_0^H$ by using (7), (8), and the Hadamard inequality [1], step (b) is derived by using (5), (7), (9), and (18) as in (29).

Then, we construct a lower bound of $C_{\min\max}$:

$$\begin{aligned}
C_{\min\max} &= \min_{|||\mathbf{H}-\mathbf{H}_0|||\le\varepsilon} \max_{\mathbf{Q}\in\mathcal{Q}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{H}\mathbf{Q}\mathbf{H}^H\right)\\
&\overset{(a)}{=} \min_{|||\mathbf{H}-\mathbf{H}_0|||\le\varepsilon} \max_{\mathbf{Q}\in\mathcal{Q}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{\Sigma}_{\mathbf{H}}\mathbf{\Lambda}_{\mathbf{Q}}\mathbf{\Sigma}_{\mathbf{H}}\right)\\
&\overset{(b)}{=} \min_{|||\mathbf{H}-\mathbf{H}_0|||\le\varepsilon} \max_{\Lambda_{\mathbf{Q}}\in\mathcal{Q}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{\Sigma}_{\mathbf{H}}\mathbf{\Lambda}_{\mathbf{Q}}\mathbf{\Sigma}_{\mathbf{H}}\right)\\
&\overset{(c)}{\ge} \min_{|||\mathbf{\Sigma}_{\mathbf{H}}-\mathbf{\Sigma}_{\mathbf{H}_0}|||\le\varepsilon} \max_{\Lambda_{\mathbf{Q}}\in\mathcal{Q}} \log\det\left(\mathbf{I}_r+\gamma\mathbf{\Sigma}_{\mathbf{H}}\mathbf{\Lambda}_{\mathbf{Q}}\mathbf{\Sigma}_{\mathbf{H}}\right)\\
&\overset{(d)}{=} \min_{\substack{\|\boldsymbol{\sigma}-\boldsymbol{\sigma}_0\|\le\varepsilon\\ \boldsymbol{\sigma}\ge\mathbf{0}}} \max_{\substack{\boldsymbol{\lambda}\in B_{\mathcal{Q}}\\ \boldsymbol{\lambda}\ge\mathbf{0}}} \sum_{i=1}^{\min\{t,r\}} \log(1+\gamma\sigma_i^2\lambda_i),
\end{aligned} \tag{31}$$

where step (a) is due to the optimal power allocation result by using (7), (8), and the Hadamard inequality [1], step (b) is due to (7), step (c) is due to the following result for unitarily invariant matrix norm: [18, Theorem 7.4.51] [21, Eq. (3.5.30)]

$$|||\mathbf{\Sigma}_{\mathbf{H}}-\mathbf{\Sigma}_{\mathbf{H}_0}||| \le |||\mathbf{H}-\mathbf{H}_0|||,$$

with $\mathbf{\Sigma}_{\mathbf{H}}$ and $\mathbf{\Sigma}_{\mathbf{H}_0}$ being the diagonal matrices in the SVDs of $\mathbf{H}$ and $\mathbf{H}_0$, and step (d) is due to (9) and (18). Then, (20) follows from (30) and (31). (19) can be proved similar to (13). By this, the theorem is proved. ■

References

[1] E. Telatar, "Capacity of multi-antenna Guassian channels," *Eur. Trans. Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, Nov. 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.: Kluwer Academic Press*, no. 6, pp. 311–335, 1998.

[3] A. Moustakas and S. Simon, "Optimizing multiple-input single-output (MISO) communication systems with general Gaussian channels: non-trivial covariance and nonzero mean," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2770–2780, Oct. 2003.

[4] S. Simon and A. Moustakas, "Optimizing MIMO antenna systems with channel covariance feedback," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 406–417, Apr. 2003.

[5] S. Jafar and A. Goldsmith, "Transmitter optimization and optimality of beamforming for multiple antenna systems," *IEEE Trans. Wirel. Commun.*, vol. 3, no. 4, pp. 1165–1175, July 2004.

[6] A. Moustakas and S. Simon, "On the outage capacity of correlated multiple-path MIMO channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3887–3903, Nov. 2007.

[7] P. Kazakopoulos, P. Mertikopoulos, A. Moustakas, and G. Caire, "Living at the edge: A large deviations approach to the outage MIMO capacity," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1984–2007, Apr. 2011.

[8] D. Blackwell, L. Breiman, and A. J. Thomasian, "The capacity of a class of channels," *Ann. Math. Statist.*, vol. 30, no. 4, pp. 1229–1241, 1959.

[9] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2632–2639, Sep. 2001.

[10] A. Goldsmith, S. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[11] S. Loyka and C. 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.

[12] S. Vorobyov, A. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[13] R. Lorenz and S. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.

[14] J. Wang and D. Palomar, "Worst-case robust MIMO transmission with imperfect channel knowledge," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3086–3100, Aug. 2009.

[15] J. Huang and A. Swindlehurst, "Robust secure transmission in MISO channels based on worst-case optimization," *IEEE Trans. Signal Process.*, vol. 60, no. 4, pp. 1696–1707, Apr. 2012.

[16] A. Wiesel, Y. Eldar, and S. Shamai, "Optimization of the MIMO compound capacity," *IEEE Trans. Wirel. Commun.*, vol. 6, no. 3, pp. 1094–1101, Mar. 2007.

[17] S. Denic, C. Charalambous, and S. Djouadi, "Information theoretic bounds for compound MIMO Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 4, pp. 1603–1617, Apr. 2009.

[18] R. A. Horn and C. R. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[19] Y. Sun, C. E. Koksal, and N. B. Shroff, "Capacity of compound MIMO Gaussian channels with additive uncertainty," 2013, Technical Report, Dept of ECE, Ohio State University. [Online]. Available: http://arxiv.org/abs/1208.4656

[20] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York: Cambridge University Press, 2011.

[21] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. New York: Cambridge University Press, 1991.

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

[23] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.