Capacity of Compound MIMO Gaussian Channels with Additive Uncertainty

Yin Sun, Member, IEEE, C. Emre Koksal, Senior Member, IEEE, and Ness B. Shroff, Fellow, IEEE.

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 matrix 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 problem has a hidden convexity property, which can be exploited to map the problem into a convex optimization problem. 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 min-max channel capacity is zero, which proves and generalizes a conjecture of Loyka and Charalambous. The key tools for showing these results are a new matrix determinant inequality and some unitarily invariant properties.

Index Terms—Channel uncertainty, compound channel, hidden convexity, multiple antenna.

I. INTRODUCTION

Multiple-input multiple-output (MIMO) techniques have been extensively used 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 matrix design.

There have 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 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, and its capacity is determined as the maximum of the worst-case mutual information (max-min channel capacity) of the set of possible channel realizations [8]. From 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 inaccurate CSI, where the channel error may be caused by estimation, interpolation, mobility, and/or feedback. In this case, the channel is typically modeled as the sum of a known nominal channel and unknown channel uncertainty. This additive channel uncertainty model has been widely utilized in studies on the fundamental limits of MIMO channels, e.g., [9]–[11], and on robust transceiver designs, 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. Arbitary 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 the 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 matrix to achieve the capacity of the compound MIMO Gaussian channel with additive channel uncertainty. We consider the case where 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 non-convex optimization problem. However, we reveal that this problem possesses a hidden convexity property, and hence can be simplified into a convex optimization problem. We first prove that the optimal transmit design is to diagonalize the nominal channel. We then show that the duality gap between

Manuscript received June 5, 2013; revised September 18, 2013; accepted September 19, 2013. This paper was presented in part at the 2013 IEEE International Symposium on Information Theory. 
Yin Sun and C. Emre Koksal are with the Department of ECE, The Ohio State University, Columbus, OH, 43210 (e-mail: sunyin02@gmail.com, koksal.2@osu.edu).
Ness B. Shroff is with the Departments of ECE and CSE, The Ohio State University, Columbus, OH, 43210 (e-mail: shroff.11@osu.edu).
Communicated by A. Lozano, Associate Editor for Communications.
Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.
II. SYSTEM MODEL

A. Notation

The following notations are used throughout the paper. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, standard upper-case letters denote sets 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. Let \( S \subset \mathbb{C}^{m \times n} \) denote sets and standard lower-case letters denote scalars. Let \( C \) denote sets and standard lower-case letters denote scalars.

The following notations are used throughout the paper. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, standard upper-case letters denote sets 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. Let \( S \subset \mathbb{C}^{m \times n} \) denote sets and standard lower-case letters denote scalars. Let \( C \) denote sets and standard lower-case letters denote scalars.

The following notations are used throughout the paper. Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, standard upper-case letters denote sets 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. Let \( S \subset \mathbb{C}^{m \times n} \) denote sets and standard lower-case letters denote scalars. Let \( C \) denote sets and standard lower-case letters denote scalars.

B. Channel Model

Consider the complex-valued Gaussian vector channel:

\[
y = \mathbf{H}x + \mathbf{n},
\]

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

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

\[
\mathbf{H} \in \mathcal{H},
\]

where \( \mathcal{H} \) is the channel region defined by

\[
\mathcal{H} \triangleq \{ \mathbf{H} : \| \mathbf{H} - \mathbf{H}_0 \|_2 \leq \varepsilon \},
\]

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

\[
\| \mathbf{A} \|_2 \triangleq \max_{\| \mathbf{x} \|_2 \leq 1} \| \mathbf{Ax} \|_2 = \max_i \sigma_i(\mathbf{A}) = \| \mathbf{A} \|_\infty.
\]

The spectral norm is a unitarily invariant matrix norm. A unitarily invariant matrix norm satisfies \([18, \text{Section 7.4.16}]\)

\[
\| \mathbf{UAV} \| = \| \mathbf{A} \| \tag{5}
\]

for all \( \mathbf{A} \in \mathbb{C}^{m \times n} \) and all unitary matrices \( \mathbf{U} \in \mathbb{C}^m \) and \( \mathbf{V} \in \mathbb{C}^n \). According to (3) and (5), the channel region \( \mathcal{H} \) is an isotropic set centered at \( \mathbf{H}_0 \), such that \( \mathbf{H}_0 + \mathbf{U}(\mathbf{H} - \mathbf{H}_0)\mathbf{V} \in \mathcal{H} \) for all \( \mathbf{H} \in \mathcal{H} \) and all unitary matrices \( \mathbf{U} \) and \( \mathbf{V} \).

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

\[
\| \mathbf{A} \|_2 \geq \| \mathbf{A} \| \tag{6}
\]

holds for all matrix \( \mathbf{A} \) and all unitarily invariant matrix norm \( \| \cdot \| \) \([18, \text{Corollary 5.6.35}]\). More discussions on the relationship among some matrix norms are provided in Section IV.

C. Power Constraint

We consider a general transmit power constraint

\[
\mathbf{Q} \in \mathcal{Q},
\]

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

\[
\mathbf{UQU}^H \in \mathcal{Q}, \tag{8}
\]

\[
\mathbf{D}(\mathbf{Q}) \in \mathcal{Q}, \tag{9}
\]

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 that a set \( \mathcal{Q} \) is unitarily invariant if it satisfies (8) and (9). One can show that each unitarily invariant \( \mathcal{Q} \) can be equivalently expressed as

\[
\mathcal{Q} = \{ \mathbf{Q} \in \mathbb{S}^t_+ : \lambda(\mathbf{Q}) \in B_\varepsilon, \lambda(\mathbf{Q}) \geq 0 \}, \tag{10}
\]

where \( B_\varepsilon \) is a nonempty compact convex set. Two typical examples of unitarily invariant power constraints are the sum power constraint \([1]\)

\[
\mathcal{Q}_1 = \{ \mathbf{Q} \in \mathbb{S}^t_+ : \text{Tr}(\mathbf{Q}) \leq P \}, \tag{11}
\]

\[
= \{ \mathbf{Q} \in \mathbb{S}^t_+ : \sum_{i=1}^t \lambda_i(\mathbf{Q}) \leq P, \lambda(\mathbf{Q}) \geq 0 \},
\]

and the maximum power constraint \([14]\)

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

III. OPTIMAL TRANSMIT COVARIANCE DESIGN

A. Main Result

The capacity of the compound MIMO Gaussian channel (1)-(3) is \([20, \text{Theorem 7.1}]\)

\[
C_{\max \min} = \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\| \mathbf{H} - \mathbf{H}_0 \|_2 \leq \varepsilon} \text{I}(\mathbf{Q}, \mathbf{H}), \tag{12}
\]

1This is different from the channel uncertainty model in \([19]\), where \( \mathbf{H} \) is within an isotropic set such that \( \mathbf{HU} \in \mathcal{H} \) for all \( \mathbf{H} \in \mathcal{H} \) and all unitary matrix \( \mathbf{U} \).
where $I(Q, H) = I(x; y)$ is the mutual information of the channel (1), i.e., [1]

$$I(Q, H) = \log \det \left( I_r + HQH^H \right).$$

Finding an efficient solution of the max-min problem (12) has been open for a long time (except in some special cases [11], [14], [16]), because $I(Q, H)$ is nonconvex with respect to $H$. However, we will show that the problem (12) possesses a hidden convexity property when (3) holds, and thus can be simplified into a convex optimization problem.

Suppose that the singular value decomposition (SVD) of the nominal channel $H_0$ is given by

$$H_0 = U_0 \Sigma H_0 V_0^H,$$

where $U_0 \in \mathbb{C}^{r \times r}$ and $V_0 \in \mathbb{C}^{t \times t}$ are unitary matrices. The first key result of this paper is stated as follows:

**Theorem 1.** If $Q$ and $H$ are nonempty sets, $H$ is defined in (3), and $Q$ satisfies the unitarily invariant properties (8) and (9), then

$$Q^* = V_0 A_Q^* V_0^H, \quad H^* = U_0 \Sigma H_0 V_0^H,$$

is a solution to Problem (12), where $U_0$ and $V_0$ are defined in (13), the diagonal matrices $A_Q^*$ and $\Sigma H_0^*$ are given by $\Sigma H_0 = \text{diag}(\Lambda^*)$ and $A_Q^* = \text{diag}(\Lambda^*)$, such that $(\sigma^*, \lambda^*)$ solves the problem

$$C_{\max \min} = \max_{A \in B_Q} \min_{\sigma \geq 0} \left\{ \min_{t,r} \sum_{i=1}^{\min\{t,r\}} \log(1 + \sigma_i^2 \lambda_i) \right\},$$

with the convex set $B_Q$ defined in (10).

**Proof:** The proof of Theorem 1 relies on the unitarily invariant properties (4), (5), (8), and (9), and a new matrix determinant inequality presented in Lemma 1 given below. The details of the proof are provided in Appendix A.

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

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

$$\min_{\|\Lambda\|_2 \leq \epsilon} \det \left[ I + (\Sigma + \Delta) A (\Sigma + \Delta)^H \right]$$

is a diagonal matrix.

**Proof:** See Appendix C.

Theorem 1 implies that the optimal transmit covariance of the MIMO Gaussian channel with worst case channel uncertainty is to diagonalize the nominal channel $H_0$. Such a solution structure was previously known only for some special cases, such as high SNR limit, low SNR limit, low rank nominal channels (rank($H_0) \leq 2$) [11], [14], [16]. In contrast, Theorem 1 holds for general nominal channels and all SNR values. Further, by Theorem 1, the problem (12) reduces to (15) with much fewer variables.

**B. The Dual Problem**

Now, we consider the dual of the max-min problem (12), which is given by the following min-max channel capacity problem

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

It is important to distinguish the capacity of the compound channel $C_{\max \min}$ and the min-max channel capacity $C_{\min \max}$: $C_{\max \min}$ can be achieved for any channel $H$ within $\mathcal{H}$, by using the same transmit covariance matrix $Q$. $C_{\min \max}$ is the minimal capacity of the channels with $H \in \mathcal{H}$, evaluating which requires knowledge of $H$ at the transmitter to obtain $Q$. We study the min-max problem (17) to gain more insight into the max-min problem (12). We consider a more general channel uncertainty region

$$\mathcal{H} = \{ H : \|H - H_0\| \leq \epsilon \},$$

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

$$\| A \| \leq \| \sigma(A) \|$$

holds for all $A \in \mathbb{C}^{m \times n}$ [21, Theorem 3.5.18]. For the special case of spectral norm, the associated vector norm in (19) is $\| \cdot \|_\infty$, as given by (4). We have the following result:

**Theorem 2.** If $Q$ and $H$ are nonempty sets, $H$ is defined in (18), and $Q$ satisfies the unitarily invariant properties (8) and (9), then

$$Q' = V_0 A_Q^* V_0^H, \quad H' = U_0 \Sigma H_0 V_0^H,$$

is a solution to Problem (17), where $U_0$ and $V_0$ are defined in (13), the diagonal matrices $A_Q^*$ and $\Sigma H_0^*$ are given by $\Sigma H_0 = \text{diag}(\Lambda^*)$ and $A_Q^* = \text{diag}(\Lambda^*)$, such that $(\sigma^*, \lambda^*)$ solves the problem

$$C_{\min \max} = \min_{\|\sigma \|_2 \leq \epsilon} \max_{\Lambda \in B_Q} \min_{\sigma \geq 0} \sum_{i=1}^{\min\{t,r\}} \log(1 + \sigma_i^2 \lambda_i),$$

with the vector norm $\| \cdot \|$ and the convex set $B_Q$ defined in (19) and (10), respectively.

**Proof:** The proof of Theorem 2 relies on the unitarily invariant properties (5), (8), (9), and (19), but not the matrix determinant inequality in Lemma 1, which has only been proved for the spectral norm case. Therefore, Theorem 2 holds for any unitarily invariant matrix norm. The details of the proof 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 $Q$ is the sum power constraint $Q_1$.

---

2 The outer optimization of $Q$ in (12) is done without knowing $H$.

3 The inner optimization of $Q$ in (17) is done with knowledge of $H$. 
C. Duality Gap is Zero

It is interesting to see that the max-min problem (12) and the min-max problem (17) have similar solution structures, as given in (14) and (20), and the difference is only in the solutions to (15) and (21). Next, we study whether (15) and (21) have a common solution for the spectral norm case.

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

\[ C_{\text{max-min}} \leq C_{\text{min-max}}. \]  

Moreover, equality holds in (22), i.e.,

\[ C_{\text{max-min}} = C_{\text{min-max}}, \]  

if and only if (15) and (21) have a common solution [22, Corollary 9.16]. It was conjectured in [11] that (23) holds for the case that \( \| \cdot \| = \| \cdot \|_2 \) and the power constraint is \( Q = Q_1 \). Here, using Theorems 1, 2, and von Neumann’s minimax theorem, we can now prove this conjecture:

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

1) The strong duality relation (23) holds.

2) Problems (15) and (21) have a common solution \( (\sigma^*, \lambda^*) \), where \( \sigma^* \) is given by

\[ \sigma_i^* = \max \{\sigma_{0,i} - \varepsilon, 0\}, \]  

and \( \lambda^* \) is determined by the convex optimization problem

\[ C_{\text{max-min}} = \max_{\lambda \in B_0} \min_{{t,r}} \sum_{i=1}^{\min \{t,r\}} \log(1 + \max \{\sigma_{0,i} - \varepsilon, 0\}^2 \lambda_i). \]  

Proof: 1) Problem (15) can be expressed as

\[ C_{\text{min-max}} = \max_{\lambda \in B_0} \min_{\lambda \in B_0} \sum_{i=1}^{\min \{t,r\}} \log(1 + \sigma_i^2 \lambda_i) \]

s.t. \( \max \{\sigma_{0,i} - \varepsilon, 0\} \leq \sigma_i \leq \sigma_{0,i} + \varepsilon, \forall i \).

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

\[ \max_{\lambda \in B_0} \sum_{i=1}^{\min \{t,r\}} \log(1 + e^{2x_i} \lambda_i) \]  

s.t. \( \log(\max \{\sigma_{0,i} - \varepsilon, 0\}) \leq x_i \leq \log(\sigma_{0,i} + \varepsilon), \forall i \),

where the objective function is concave in \( \lambda \) and convex in \( x \) [23]. Similarly, (21) can be reformulated as the following convex optimization problem:

\[ \min_{\lambda \in B_0} \left\{ \max_{\lambda \in B_0} \sum_{i=1}^{\min \{t,r\}} \log(1 + e^{2x_i} \lambda_i) \right\} \]

s.t. \( \log(\max \{\sigma_{0,i} - \varepsilon, 0\}) \leq x_i \leq \log(\sigma_{0,i} + \varepsilon), \forall i \).

Let us use \( f(\lambda, x) \) to denote the objective function in (26) and (27). We say a point \( (\lambda_0, x_0) \) is a saddle point of \( f \) if

\[ f(\lambda_0, x_0) = \min_{\sigma_{0,i} - \varepsilon} \max_{\sigma_{0,i} + \varepsilon} f(\lambda_0, x) = \max_{\lambda \in B_0} \min_{\lambda \in B_0} f(\lambda, x_0). \]  

By von Neumann’s minimax theorem [22, Theorem 9.D], we have: (1) the function \( f \) has a saddle point; (2) a point \( (\lambda_0, x_0) \) is a saddle point of \( f \) if and only if \( (\lambda_0, x_0) \) is a common solution to (26) and (27). Therefore, (26) and (27) must have the same optimal objective value, and (23) follows.

2) Let us define \( x^* \) as \( x_i^* = \log(\max \{\sigma_{0,i} - \varepsilon, 0\}) \). By von Neumann’s minimax theorem, we only need to show that \( (x^*, \lambda^*) \) is a saddle point of \( f \). When \( (\lambda_0, x_0) \) is replaced by \( (x^*, \lambda^*) \), the minimization problem in (28) can be separated into several subproblems, i.e.,

\[ \min_{x_i} \log(1 + e^{2x_i} \lambda_i^*) \]

s.t. \( \log(\max \{\sigma_{0,i} - \varepsilon, 0\}) \leq x_i \leq \log(\sigma_{0,i} + \varepsilon), \forall i \),

and the solution is given by \( x_i^* \). On the other hand, according to (25), \( \lambda^* \) is the solution to the maximization problem in (28). Therefore, \( (x^*, \lambda^*) \) is a saddle point of \( f \), and thus also a common solution of (15) and (21).

IV. DISCUSSION

It is known that any matrix norm can be bounded within a constant multiple of the spectral norm: for any finite \( t, r \) and any matrix norm \( \| \cdot \|, \) there exist \( \alpha_1, \alpha_2 > 0 \) such that

\[ \alpha_1 \| H - H_0 \|_2 \leq \| H - H_0 \| \leq \alpha_2 \| H - H_0 \|_2, \]  

where \( \alpha_1 \) and \( \alpha_2 \) are irrelevant to the channel error \( H - H_0 \) but depend on \( \| \cdot \| \) [18]. The values of \( \alpha_1 \) and \( \alpha_2 \) are summarized in Table I of [11] for some popular matrix norms. In particular, when \( \| \cdot \| \) is a unitarily invariant matrix norm, \( \alpha_2 = 1 \) as in (6). Using this and our results in Theorem 1 and 3, we have shown that the covariance design problem (12) is a convex optimization problem in nature, if the channel uncertainty region \( \mathcal{H} \) is characterized by the spectral norm.

Another interesting open problem is whether the matrix determinant inequality in Lemma 1 holds for some other unitarily invariant matrix norms, e.g., the Frobenius (Euclidean) norm

\[ \| A \|_F \triangleq \sqrt{\sum_{i,j} |a_{ij}|^2} = \| \sigma(A) \|_2. \]  

If one can generalize Lemma 1 to another unitarily invariant matrix norm, then Theorems 1 and 3 also hold for this particular matrix norm, with the infinite norm \( \| \cdot \|_\infty \) in (15) replaced by the corresponding vector norm defined by (19).

We have found a counterexample, which shows that Lemma 1 does not hold for the matrix norm \( \| \cdot \| = \| \sigma(\cdot) \|_1 \). Consider the matrices

\[ \Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}. \]  

If \( \Delta \) is restricted as a diagonal matrix, one can numerically calculate that the optimal objective value of the problem

\[ \min_{\| \sigma(\Delta) \|_1 \leq 1} \det [I + (\Sigma + \Delta)A(\Sigma + \Delta)^H] \]  

is

\[ \log(4). \]
is 15.63. However, when

\[
\Delta = \begin{bmatrix} -0.5 & -0.5 \\ -0.5 & -0.5 \end{bmatrix},
\]

(33)

the value of the objective function in (32) is 15.5. Therefore, the solution to (31) and (32) is not necessarily a diagonal matrix. For other unitarily invariant matrix norms, we still do not know whether Lemma 1 holds or not.

V. 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 closed-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 min-max channel capacity. When the channel uncertainty region is formulated by other matrix norms, the capacity of the compound MIMO channel remains open.

APPENDIX A

PROOF OF THEOREM 1

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

\[
\begin{align*}
C_{\text{max min}} & \leq \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{H} \mathbf{Q} \mathbf{H}^H) \\
& = \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{Q} \mathbf{V}_0^H \mathbf{\Sigma}_H^H),
\end{align*}
\]

(34)

(35)

where \( \mathbf{U}_0 \) and \( \mathbf{V}_0 \) are defined in (13), (34) is due to the additional constraint in the inner minimization problem, and (35) is due to \( \mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H \) and \( \det (\mathbf{I} + \mathbf{A} \mathbf{B}) = \det (\mathbf{I} + \mathbf{B} \mathbf{A}) \).

Let us define \( \mathbf{Q} \triangleq \mathbf{V}_0^H \mathbf{Q} \mathbf{V}_0 \) and use \( \mathcal{D}(\mathbf{Q}) \) to denote the diagonal matrix that has the same diagonal elements as \( \mathbf{Q} \). We then attain

\[
\begin{align*}
C_{\text{max min}} & \leq \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{Q} \mathbf{\Sigma}_H^H) \\
& = \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{Q} \mathbf{\Sigma}_H^H) \\
& = \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{D}(\mathbf{Q}) \mathbf{\Sigma}_H^H) \\
& \leq \max_{\mathbf{D}(\mathbf{Q}) \in \mathcal{Q}} \min_{\mathbf{\Sigma}_H \mathbf{H} \mathbf{D}(\mathbf{Q})^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{D}(\mathbf{Q}) \mathbf{\Sigma}_H^H) \\
& = \max_{\lambda \in \mathcal{B}(\mathbf{Q})} \min_{\sigma - \sigma_0 \geq \varepsilon} \log \det \left( \mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{D}(\mathbf{Q})^H \mathbf{\Sigma}_H^H \right) \min_{\lambda \geq 0} \min_{\sigma - \sigma_0 \geq \varepsilon} \sum_{i=1}^n \log(1 + \sigma_i^2 \lambda_i),
\end{align*}
\]

(36)

(37)

(38)

(39)

where (36) is derived from (8), (37) is due to \( \| \mathbf{\Sigma}_H - \mathbf{\Sigma}_H^H \|_2 = \| \mathbf{H} - \mathbf{H}_0 \|_2 \) which is derived from \( \mathbf{H}_0 = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H \), \( \mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H \) and (5), (38) is due to the Hadamard inequality \( \det(\mathbf{A}) \leq \prod_i A_{ii} \) (39) is because the feasible region \( \mathcal{D}(\mathbf{Q}) \in \mathcal{Q} \) is larger than the region \( \mathbf{Q} \in \mathcal{Q} \) according to (9), and (40) is due to (4) and (10) with \( \lambda_i \) representing the diagonal entries of \( \mathbf{D}(\mathbf{Q}) \).

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

\[
\begin{align*}
C_{\text{max min}} & \geq \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{Q} \mathbf{\Sigma}_H^H) \\
& = \max_{\mathbf{Q} \in \mathcal{Q}} \min_{\mathbf{H} = \mathbf{U}_0 \mathbf{\Sigma}_H \mathbf{V}_0^H} \log \det (\mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{D}(\mathbf{Q}) \mathbf{\Sigma}_H^H) \\
& \geq \max_{\mathbf{Q} \in \mathcal{Q}} \log \det \left( \mathbf{I}_r + \mathbf{\Sigma}_H \mathbf{D}(\mathbf{Q})^H \mathbf{\Sigma}_H^H \right) \min_{\lambda \geq 0} \min_{\sigma - \sigma_0 \geq \varepsilon} \sum_{i=1}^n \log(1 + \sigma_i^2 \lambda_i),
\end{align*}
\]

(40)

(41)

(42)

(43)

(44)

(45)

(46)

where (44) we have used (5) and (8), (45) is due to (4) and (10).
APPENDIX B
PROOF OF THEOREM 2

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

$$C_{\text{min-max}} \leq \min_{H \in H, H \neq H_0} \max_{Q \in \cal Q} \log \det (I_r + HQH^H)$$

where \(H_0\) is due to the fact that the optimal power allocation result is of the form $Q = V_0 \Lambda_Q V_0^H$ by using (8), (9), and the Hadamard inequality \[1\], (48) is derived by using (5), (8), (10), and (19) as in (46).

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

$$C_{\text{min-max}} = \min_{H \neq H_0} \max_{Q \in \cal Q} \log \det (I_r + HQH^H)$$

where (49) is due to the optimal power allocation result by using (8), (9), and the Hadamard inequality \[1\], (50) is due to (8), (51) is due to the following result for unitarily invariant matrix norm: \[18\, \text{Theorem} \ 7.4.51 \] \[21\, \text{Eq.} \ (3.5.30)\]

$$\|\Sigma_H - \Sigma_{H_0}\| \leq \|H - H_0\|,$$

with $\Sigma_H$ and $\Sigma_{H_0}$ being the diagonal matrices in the SVDs of $H$ and $H_0$, and (52) is derived from (10) and (19). Combining (48) and (52), the optimal objective value is given by (21). Further, similarly to (46), we can show that (20) is an optimal solution of (17). Hence, the theorem is proven. \[\blacksquare\]

APPENDIX C
PROOF OF LEMMA 1

In order to prove Lemma 1, we first need to show the following result:

Lemma 2. Let $\Sigma = [\Sigma_{ii}] \in \mathbb{R}^{r \times t}$ be a diagonal matrix with non-negative diagonal entries $\Sigma_{ii} \geq 0$ for $i = 1, \ldots, \min\{r, t\}$, and $\Delta \in \mathbb{C}^{r \times t}$ be a matrix satisfying $\|\Delta\|_2 \leq \varepsilon$.

1) The following inequality holds:

$$\det \left( [\Sigma + \Delta]^H (\Sigma + \Delta) \right) \geq \prod_{j=1}^t \max \{\Sigma_{jj} - \varepsilon, 0\}^2,$$

where $\Sigma_{jj}$ are defined by $\Sigma_{jj} = 0$ for $\min\{r, t\} < j \leq t$.

2) Let $S$ be a proper subset of $\{1, 2, \ldots, t\}$, then

$$\det \left\{ [\Sigma(S) + \Delta(S)]^H [\Sigma(S) + \Delta(S)] \right\} \geq \prod_{j \in S} \max \{\Sigma_{jj} - \varepsilon, 0\}^2,$$

where $\Sigma(S)$ denotes the submatrix of $\Sigma$ obtained by deleting the rows and columns complementary to those indicated by $S$ from $\Sigma$.

Proof: 1) It is known that for any $A, B \in \mathbb{C}^{r \times t}$ the following singular value inequality holds \[24\, \text{Eq.} \ (5.12.15)\], \[18\, \text{Corollary} \ 7.3.8\]

$$|\sigma_i(A + B) - \sigma_i(A)| \leq \|B\|_2, \quad \forall i = 1, \ldots, \min\{t, r\}.$$  

By this, we have

$$\sigma_i(\Sigma + \Delta) \geq \max \{\sigma_i(\Sigma) - \|\Delta\|_2, 0\}$$

$$\geq \max \{\sigma_i(\Sigma) - \varepsilon, 0\}, \quad \forall i = 1, \ldots, \min\{t, r\},$$

where the maximization is due to the fact $\sigma_i(\Sigma + \Delta) \geq 0$. Moreover, $\sigma_i(\Sigma + \Delta) = 0$ for $\min\{t, r\} < i \leq t$. Hence, we attain

$$\det \left( [\Sigma + \Delta]^H (\Sigma + \Delta) \right) = \prod_{j=1}^t \sigma_j(\Sigma + \Delta)^2$$

$$\geq \prod_{j=1}^t \max \{\sigma_j(\Sigma) - \varepsilon, 0\}^2$$

$$= \prod_{j=1}^t \max \{\Sigma_{jj} - \varepsilon, 0\}^2.$$

2) Since $\Sigma$ is a diagonal matrix, after deleting the rows and columns, the singular values of the submatrix $\Sigma(S)$ are given by $\{\sigma_{ii} : i \in S\}$. Moreover, after deleting some rows and columns, the spectral norm of $\Delta(S)$ satisfies $\|\Delta(S)\|_2 \leq \|\Delta\|_2$ \[18\, \text{Theorem} \ 7.3.9\]. Therefore

$$\det \left\{ [\Sigma(S) + \Delta(S)]^H [\Sigma(S) + \Delta(S)] \right\}$$

$$= \prod_{j=1}^{\|S\|} \sigma_j(\Sigma(S) + \Delta(S))^2$$

$$\geq \prod_{j=1}^{\|S\|} \max \{\sigma_j(\Sigma(S)) - \|\Delta(S)\|_2, 0\}^2$$

$$\geq \prod_{j=1}^{\|S\|} \max \{\sigma_j(\Sigma(S)) - \varepsilon, 0\}^2$$

$$= \prod_{j \in S} \max \{\Sigma_{jj} - \varepsilon, 0\}^2.$$
3) For any given $S \subseteq \{1, 2, \ldots, t\}$, one can interchange the rows and columns of $A$ and $(\Sigma + \Delta)$ by multiplying with two permutation matrices $P \in S_t^+$ and $Q \in S_t^+$ as $B = PA^T P$ and $\Phi = Q(\Sigma + \Delta)P$, such that $A(S)$ and $(\Sigma(S) + \Delta(S))$ are the leading submatrices of $B$ and $\Phi$, respectively, i.e.,

$$B = \begin{pmatrix} A(S) & C \\ C^T & D \end{pmatrix}, \quad \Phi = \begin{pmatrix} \Sigma(S) + \Delta(S) & M \\ N & G \end{pmatrix},$$

Since $P^H = P$, $Q^H = Q$, $P^2 = I$ and $Q^2 = I$, we attain

$$B = P(\Sigma + \Delta)^H QQ(\Sigma + \Delta)P = \Phi^H \Phi,$$

thereby

$$A(S) = (\Sigma(S) + \Delta(S))^H (\Sigma(S) + \Delta(S)) + N^H N,$$

Then,

$$A(S) \geq (\Sigma(S) + \Delta(S))^H (\Sigma(S) + \Delta(S)) \geq 0,$$

which further implies that [18, Corollary 7.7.4]

$$\det[A(S)] \geq \det[(\Sigma(S) + \Delta(S))^H (\Sigma(S) + \Delta(S))].$$

Finally, by using (54), the result in (55) follows.

Using part (3) of Lemma 2, we can establish the following lemma:

**Lemma 3.** Let $\Sigma = [\varsigma_{ij}] \in \mathbb{R}^{r \times t}$ and $D = [d_{ij}] \in \mathbb{R}^{r \times t}$ be two diagonal matrices with non-negative diagonal entries, and $\Delta \in \mathbb{C}^{r \times t}$ be a matrix satisfying $\|\Delta\|_2 \leq \varepsilon$. Then

$$\det[I_t + (\Sigma + \Delta)^H (\Sigma + \Delta)D] \geq \prod_{j=1}^t (1 + \max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}),$$

(57)

where equality is achieved by the diagonal matrix $\Delta^*$ with diagonal entries given by

$$\Delta^*_{jj} = \varsigma_{jj} - \max\{\varsigma_{jj} - \varepsilon, 0\}, \quad j = 1, \ldots, \min\{t, r\}.$$

Then, Lemma 1 follows from (57).

*Proof:* We prove the inequality

$$\det[I_t + (\Sigma + \Delta)^H (\Sigma + \Delta)D] \geq \prod_{j=1}^t (1 + \max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj})$$

(58)

by induction, where $\varsigma_{jj}$ for $\min\{t, r\} < j \leq t$ are defined as $\varsigma_{jj} = 0$.

Let $B = (\Sigma + \Delta)^H (\Sigma + \Delta)D$. By part (3) of Lemma 2 and $D$ being a diagonal matrix, we attain

$$\det[B(S)] \geq \prod_{j=1}^t (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}),$$

(59)

For any set $S$ satisfying $\{1\} \subseteq S \subseteq \{1, 2, \ldots, t\}$, by the cofactor (Laplace) expansion [24, Eq. (6.2.5)] of $\det[B(S)]$ and (59), we attain

$$\det[E_{11} + B(S)] = \det[B(S)] + \det[B(S \setminus \{1\}]) \geq \prod_{j=1}^t (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}) + \prod_{j=1}^t (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}).$$

(60)

Similarly, for any set $S$ satisfying $\{1, 2\} \subseteq S \subseteq \{1, 2, \ldots, t\}$, by the cofactor expansion of $\det[E_{11} + B(S)]$ and (60), we have

$$\det[E_{11} + E_{22} + B(S)] = \det[E_{11} + B(S)] + \det[E_{11} + B(S \setminus \{2\})]$$

$$\geq \prod_{j=1}^t (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}) + \prod_{j=1}^t (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}).$$

(61)

Suppose that for any set $S$ satisfying $\{1, 2, \ldots, k\} \subseteq S \subseteq \{1, 2, \ldots, t\}$, the following inequalities hold

$$\det[E_{11} + \cdots + E_{kk} + B(S)] \geq \prod_{j=1}^k (1 + \max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}) \prod_{j \in S \setminus \{1, 2, \ldots, k\}} (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}).$$

Then, for any set $S$ satisfying $\{1, 2, \ldots, k, \ldots, k + 1\} \subseteq S \subseteq \{1, 2, \ldots, t\}$, by the cofactor expansion of $\det[E_{11} + \cdots + E_{kk} + B(S)]$ and (61), we have

$$\det[E_{11} + \cdots + E_{kk} + E_{k+1,k+1} + B(S)]$$

$$\geq \prod_{j=1}^k (1 + \max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}) \prod_{j \in S \setminus \{1, 2, \ldots, k\}} (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj})$$

$$+ \prod_{j=1}^k (1 + \max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}) \prod_{j \in S \setminus \{1, 2, \ldots, k+1\}} (\max\{\varsigma_{jj} - \varepsilon, 0\}^2 d_{jj}).$$

(62)

By induction, the result of (58) follows.

Finally, (58) reduces to (57) since $\varsigma_{jj} = 0$ for $\min\{t, r\} < j \leq t$. ■
REFERENCES


