**WORST CASE ROBUST DOWNLINK BEAMFORMING ON THE RIEMANNIAN MANIFOLD**

*Dana Ciochina, Marius Pesavento*  
TU Darmstadt  
Communications Systems  
D-64283 Darmstadt Germany

**Kon Max Wong**  
McMaster University  
Dept of Electrical and Computer Engineering  
Ontario, Canada

**ABSTRACT**
In this paper we consider a new perspective on the worst case robust multiuser downlink beamforming problem with imperfect second order channel state information at the transmitter. Recognizing that all channel covariance matrices form a Riemannian manifold, we propose to use a measure properly defined along this manifold in order to model the set of mismatched channel covariance matrices for which robustness shall be guaranteed. This leads to a new robust beamforming problem formulation for which a convex approximation is derived. Simulation results show a dramatically improved performance of the proposed scheme, both in terms of transmitted power and constraint satisfaction, as compared to the previous methods.

**Index Terms**— robust downlink beamforming, imperfect CSI, Riemannian manifold

1. INTRODUCTION

When multiple antennas at a basestation (BS) are available, beamforming techniques can be used to ensure a desired quality of service (QoS) to the users in the network [1]. However, the performance of beamforming techniques substantially depends on the quality of the channel state information (CSI) available at the transmitter. Since the errors in the CSI are inherent due to, e.g., erroneous CSI estimation [2], quantization and feedback delay [3], the design of robust beamformers that consider the channel imperfections is of major practical importance.

A common approach to ensure robustness in this context, is to construct beamformers that satisfy the QoS constraints for all possible channel states within a region suitably defined around the presumed CSI. When instantaneous CSI is available at the transmitter, robust beamformer designs have been proposed in [1]-[6], in which the Frobenius norm has been considered as a measure to limit the uncertainty in the actual covariance matrix but also in its matrix decomposition to simplify the positive semidefiniteness constraints that apply. A different approach has been taken in [15] where trace bounds are used, which can be proven to be more suitable to characterize specific types of errors as e.g., errors occurring due to finite sampling effects in the CSI estimation.

Our proposed approach is motivated by the observation that when second order statistical CSI is available at the transmitter, the region in which the mismatched CSI must be considered consists of positive definite matrices which form a Riemannian manifold [16]-[18]. Thus in order to correctly characterize this set, proper Riemannian distances are more appropriate than the previously used Frobenius norms. Therefore we formulate a new worst case robust beamforming problem, in which we use the Riemannian distance previously derived in [17] to define the area around the presumed CSI, for which the signal-to-interference-and-noise (SINR) constraints are guaranteed. Using Lagrange duality theory the new beamforming problem is reduced to a convex approximation. Simulations show that our new design significantly outperforms the existing methods in terms of constraint satisfaction and transmitted power. While in this paper our robust beamforming design is applied to a multiple-input single output scenario, the ideas presented can be easily extended to more complicated scenarios including, e.g., multigroup multicasting [15] or cognitive radio [19], [20].

**Notation:** Throughout this paper we use $\text{Tr}\{\cdot\}, \mathbb{E}\{\cdot\}$ to denote the trace and expectation of a square matrix, $(\cdot)^+, \text{vec}(\cdot)$, $\|\cdot\|_F$ and $\mathcal{R}(\cdot)$, represent the pseudoinverse, vectorization, Frobenius norm and the range of the matrix, while $\otimes$, $\mathbf{I}_N$ and $\mathcal{U}$ stand for the Kronecker product, the identity matrix of dimension $N$ and the set of unitary matrices respectively.

2. SYSTEM MODEL

Consider a BS with $N$ transmit antennas serving $K$ single-antenna users. The signal transmitted at the BS is given by $x(n) = \sum_{k=1}^{K} \sqrt{p_k} u_k s_k(n)$, where $p_k, u_k, s_k(n)$ denote the power, the unit norm beamforming weight vector and the zero mean unit norm variance information symbol transmitted to
the $k$th user at time instant $n$, respectively. The signal at the $k$th receiver can be written as
\[ y_k(n) = h_k^T x(n) + z_k(n), \quad k = 1, \ldots, K \] (1)
where $h_k \triangleq [h_{k,1}, \ldots, h_{k,N}]^T$ denotes the downlink channel vector of the $k$th user, with $h_{k,m}$ being the channel coefficient between the $m$th transmit antenna and the $k$th receiver and $z_k(n)$ is the additive complex circular Gaussian noise of zero mean and variance $\sigma_k^2$. At the transmitter, the available estimates of the channel covariance matrices can be written as $\hat{R}_k = R_k - \Delta_k$, where $R_k = E \{ h_k h_k^H \}$ and $\Delta_k$ represent the true covariance matrix and the estimation error for the $k$th user respectively.

The aim of the robust downlink beamforming problem is to design beamformers which minimize the total transmitted power while satisfying imposed SINR targets for all scheduled users in the network and for all possible mismatch matrices which lie within a predefined distance from the available channel covariance matrices. For a properly defined distance function $d$ and using $d_k \triangleq d(\hat{R}_k + \Delta_k, \hat{R}_k)$ to specify the distance between the true and estimated channel covariance matrices of user $k$, we write the worst case robust downlink beamforming as
\[
\min_{\{w_i\}} \sum_{i=1}^K \|w_i\|^2 \\
\text{s.t.} \quad \frac{w_i^H}{\sqrt{\sigma_k^2+k}} \left( \hat{R}_k + \Delta_k \right) w_i \geq \gamma_k, \quad (2b)
\]
where $\gamma_k$ and $\alpha_k$ are the imposed SINR threshold and error bound for the $k$th user.

2.1. Measure Proposed to Characterize the Uncertainty Region

In order to characterize the uncertainty region for the channel mismatches, a proper distance must be introduced to measure the dissimilarity between the true and the estimated CSI. Since the Frobenius norm is known to be the shortest distance between two points in the Euclidean space, it has been considered as a reasonable choice for modelling certain uncertainty sets around the presumed CSI. Indeed, the use of this norm is well justified when limiting the mismatches in the estimates of the instantaneous channels $h_k$, available at the transmitter. This is because in this case the errors may occur arbitrarily inside a bounded set. This is however not the case when the available CSI is based on second order statistics. Due to their positive semidefiniteness property, the mismatched covariance matrices cannot be considered as free points in the Euclidean space. They form instead a Riemannian manifold, in which distances are not correctly characterized by Frobenius norms, but by properly defined Riemannian measures. Riemannian distances have been derived and successfully used in different signal processing applications, e.g., in signal classification and feature detection [17], [18]. In this paper, in order to measure the dissimilarity between the true and mismatched covariance matrices on the Riemannian manifold, we adopt the Riemannian distance as derived in [17]:
\[
d_R(\hat{R}_k, \hat{R}_k + \Delta_k) = \sqrt{2 Tr \{ \hat{R}_k \} + Tr \{ \Delta_k \} - 2 Tr \left\{ \left( \hat{R}_k^{1/2} (\hat{R}_k + \Delta_k) \hat{R}_k^{1/2} \right)^{1/2} \right\}}. \quad (3)
\]
Defining $A_k \triangleq \gamma_k \sum_{i=1}^K w_i w_i^H - \hat{R}_k$, and employing (3) the optimization sub-problem in (2b) can be written as
\[
\min_{\Delta_k} \{ \text{Tr} \{ \Delta_k A_k \} + \text{Tr} \{ \hat{R}_k A_k \} + \sigma_k^2 \gamma_k \} \geq 0 \quad (4a)
\]
s.t. $d_R^2(\hat{R}_k, \hat{R}_k + \Delta_k) \leq \alpha_k^2 \hat{R}_k \quad (4c)
\]
where $\alpha_{R,k}$ is the $k$th bound on the Riemannian distance. Therefore, our proposed robust beamforming problem becomes:
\[
\min_{\{w_i\}} \sum_{i=1}^K \|w_i\|^2 \quad \text{s.t.} \quad (4) \text{is satisfied for } k = 1 \ldots K. \quad (5)
\]

3. PROPOSED BEAMFORMING APPROACH

The problem in (5) is generally non-convex and therefore difficult to solve. In this section we derive a convex approximation of (5). To this aim, we rewrite the optimization sub-problems (4) as closed form expressions, which can be reduced to convex reformulations. Due to the complicated form of the expressions (5), we first derive a simple approximation for the Riemannian distance, which we use in the following analysis.

Lemma 1: Let $M_1$ and $M_2$ be two positive definite Hermitian matrices. Then
\[
\text{Tr} \left\{ \left( M_1^{1/2} M_2 M_1^{1/2} \right)^{1/2} \right\} \geq \text{Tr} \left\{ M_1^{1/2} M_2^{1/2} \right\} \quad (6)
\]
Proof: The term on the left hand side of (6) is the optimum of
\[
\max_{U_1, U_2 \in \mathcal{U}} \text{Re} \left\{ \text{Tr} \left\{ U_1 U_2^H M_1^{1/2} M_2 M_1^{1/2} \right\} \right\}, \quad (7)
\]
which is attained when $U_1$ and $U_2$ are right and left singular matrices of $M_1^{1/2} M_2^{1/2}$ [32]. Since $U_1 = U_2 = I_N$ are also feasible solutions of (7), the inequality in Lemma 1 holds.

Applying the result of Lemma 1 on the expression in (5), it immediately follows that
\[
d_R^2(\hat{R}_k, \hat{R}_k + \Delta_k) \leq \text{Tr} \left\{ \left( \hat{R}_k + \Delta_k \right)^{1/2} - \hat{R}_k \right\}^2 \quad (8)
\]
Using the upper bounds in (8), we can strengthen (4b) in (4) and obtain the inner approximations
\[
\min_{\Delta_k} \{ \text{Tr} \{ \Delta_k A_k \} + \text{Tr} \{ \hat{R}_k A_k \} + \sigma_k^2 \gamma_k \} \leq \alpha_k^2 \hat{R}_k, \quad (9c)
\]
where (4c) is satisfied.
whose optimal value is denoted by $\Upsilon_k^\lambda$. We have all elements to make the following statement.

**Proposition 1:** A sufficient condition for the worst case SINR constraints in (4) to be satisfied, i.e., $\Upsilon_k^\lambda \geq 0$ is that there exists a set of non-negative $\lambda_k$ such that

$$
X_k \triangleq \left( -I_N \otimes A_k + \lambda_k I_{N^2} b_k(\lambda_k) \right) \geq 0 \quad (10)
$$

where

$$
b_k(\lambda_k) = -\lambda_k \text{vec}(R_k^\lambda) \quad (11a)
$$

$$
c_k(\lambda_k) = \lambda_k \text{Tr}\{\hat{R}_k\} - \sigma_k^2 - \lambda_k \text{Tr}_{R_k} \quad (11b)
$$

Proof: Introducing $Q_k \triangleq \left( \hat{R}_k + \Delta_k \right)^{1/2}$ problem (9) becomes equivalent to

$$
\begin{align}
\min_{Q_k} - \text{Tr}\left\{Q_k A_k Q_k^H\right\} - \sigma_k^2
\quad \text{s.t.} \quad \left\{Q_k - \hat{R}_k^{1/2} \left(Q_k - \hat{R}_k^{1/2}\right)^H\right\} \leq \alpha_k^2 \quad (12a)
\end{align}
$$

$$
Q_k = Q_k^H \quad (12b)
$$

$$
\text{Note, that in (12), the constraints (12c) are redundant since in their absence any optimal solution of (12) is still Hermitian. This can be proven as follows. If $Q_k^\lambda$ is optimal for the problem formed by (12a) and (12b), then $Q_k^{\lambda H}$ is also optimal and the first-order optimality conditions (21) imply that}
$$
Q_k^\lambda \left(-A_k + \lambda_k I_N\right) = \left(-A_k + \lambda_k I_N\right) Q_k^\lambda = \hat{R}_k^{1/2} \quad (13)
$$

From (13) and further using (22) Th 1.3.12 it follows that $Q_k^\lambda = Q_k^H$. Therefore, with $q_k = \text{vec}(Q_k)$, (13) can be equivalently written as

$$
\begin{align}
\min_{q_k} - q_k^H \left(I_N \otimes A_k\right) q_k - \sigma_k^2
\quad \text{s.t.} \quad q_k^H q_k - 2\text{Re}\left\{\text{vec}^H\left(\hat{R}_k^{1/2}\right) q_k\right\} + \text{Tr}\{\hat{R}_k\} - \alpha_k^2 \leq 0
\end{align}
$$

Since (14) represents a quadratic program with a single quadratic constraint, strong duality holds if it admits a strictly feasible solution. In the case of (14), strong duality can be claimed for any positive $\alpha_k$, due to the strict feasibility of $q_k = \text{vec}(\hat{R}_k^{1/2})$.

Using the notations in (11a) and (11b), the dual problem of (14) can be written as

$$
\begin{align}
\max_{\lambda_k} c_k(\lambda_k) - b_k^H(\lambda_k) \left(-I_N \otimes A_k + \lambda_k I_{N^2}\right)^+ b_k(\lambda_k) \quad (15)
\quad \text{s.t.} - I_N \otimes A_k + \lambda_k I_{N^2} \geq 0
\end{align}
$$

$$
b_k(\lambda_k) \in R(-I_N \otimes A_k + \lambda_k I_{N^2})
$$

Due to the strong duality property, it follows that, in order to satisfy the worst case SINR constraints, it is sufficient to find a $\lambda_k$ for which a non-negative objective value for the dual problem (15) is obtained. Furthermore, since the objective of the dual is always smaller or equal than the primal for all feasible points, it is sufficient to find one $\lambda_k$ in the feasible set of (15), such that the objective is non-negative.

Finally, using Schur complement, it is proven in (21) Appendix A5.4 that $\lambda_k$ is in the feasible set of (15) and achieves a non-negative objective value if and only if $X_k \geq 0$, where $X_k$ were defined in equation (10).

![Fig. 1. Transmitted power for uniform error on the covariance matrix](image)

From Proposition 1, it follows that replacing in (5) the worst case SINR constraints (4) by (10) guarantees that the former are satisfied. The advantage of the latter constraints is however, that they can be converted into a convex reformulation, as seen in the following. Introducing $W_k \triangleq w_k w_k^H$, for $k = 1, \ldots, K$ and using the semidefinite relaxation procedure (23), the initial problem (5) can be rewritten as

$$
\begin{align}
\min_{\{W_i, \lambda_i\}} \sum_{k=1}^{K} \text{Tr}\{W_i\}
\quad \text{s.t.} \quad (10) \text{ is satisfied, } W_k \succeq 0, \lambda_k \geq 0, k = 1, \ldots, K
\end{align}
$$

If the resulting optimal matrices $W_k$ exhibit a rank larger than one randomization techniques (23) can be used.

4. SIMULATIONS

In our simulations we consider two scenarios in which different error models are assumed for the covariance matrices. We show the performance of our algorithm in terms of transmitted power and SINR satisfaction as compared to the methods in [12] and [14].

In order to illustrate the choice of the thresholds we use in our comparison, we briefly present the measures employed in [12] and [14] to characterize the mismatches. In (12) the uncertainty region is bounded as $\|\Delta_k\|_F \leq \alpha_{F,k}$ where $\alpha_{F,k}$ denotes the imposed uncertainty thresholds under Frobenius norm. In (14), in order to simplify the positive semidefiniteness constraints on the $\hat{R}_k + \Delta_k$, the decomposition $\hat{R}_k + \Delta_k = \left(P_k + \Delta_k\right) \left(P_k + \Delta_k\right)^H$ is considered, where $P_k$ are such that $\hat{R}_k = P_k P_k^H$ and $\Delta_k$ can be written with respect to $P_k$ and $\Delta_k$. Then, Frobenius norm is used to bound both the mismatch on the available covariance matrix $\Delta_k$ and the mismatch on the square root of the erroneous CSI $\Delta_k^H$. Denoting the latter threshold by $\eta_k$, it can be proven that it relates to $\alpha_{F,k}$ as (14)

$$
\alpha_{F,k} \leq 2\eta_k \|P_k\|_F + \eta_k^2 \quad (17)
$$

For our first simulation we use the scenario in (14), where the errors are generated uniformly within a sphere around the true covariance matrix. This reflects the case when quantized covariance based CSI is available at the transmitter.
the histogram of the normalized QoS defined as
\[ \eta_k = 1 / \sum_{i=1}^{N_k} \| \hat{s}_i - \mu \|^2 \] an SINR level of 4dB, and thresholds \( \alpha \) where the angular spread is

covariance matrix

Fig. 2

The true covariance matrices are modelled as in [24], where the angular spread is 20° and users are positioned at \( 10^\circ, 10^\circ + \theta_s, 10^\circ + 2\theta_s \) with \( \theta_s \) the separation angle. For an SINR level of 4dB, and thresholds \( \alpha_{F,k} = 0.15 \) and \( \alpha_{R,k} \) as well as \( \eta_k \) chosen according to [17], the transmitted power is plotted in Figure 1. For further comparison, we plot in Figure 2 the histogram of the normalized QoS defined as

\[ \frac{w_k^H (\hat{R}_k + \Delta_k) w_k}{\gamma_k} \sum_{i \neq k}^{K} w_i^H (\hat{R}_k + \Delta_k) w_i + \gamma_k \sigma_k^2 \] (18)

We note from Figure 2 that our method respects all constraints without oversatisfying them as much as the previous techniques in [12] and [14]. This further confirms the better performance obtained by our method in terms of transmitted power.

In the second scenario we assume that the channels at the receiver are estimated using a Least Squares (LS) algorithm. We consider the number of training symbols \( N \), the training power 15dB and the variance of the estimation error –20dB. The estimated covariance matrices are constructed based on \( N_s = 512 \) training snapshots. In a first stage, in order to obtain more realistic bounds we consider the true covariance matrices \( \hat{R}_k \) as defined in [24] and realize a realization of the true channel as \( \hat{h}_k = U_k \Lambda_k^{1/2} v_k \), where \( U_k \) and \( \Lambda_k \) come from the eigenvalue decomposition \( \hat{R}_k = U_k \Lambda_k U_k^H \) and \( v_k \) are random Gaussian vectors. This corresponds to the general channel model with correlated fading. The LS channel estimate in this case is given by \( \hat{h}_k = h + T^* e_k \), where \( T \) is the training matrix that can, e.g., be chosen as

Fig. 3

a weighted DFT and \( e \) is the zero mean complex circular Gaussian estimation error. Then the estimated covariance matrix is \( \hat{R}_k = 1/N_s \sum_{i=1}^{N_k} h_i^H h_i \). For 3000 true and estimated covariance matrices generated according to the procedure described above the Frobenius and Riemannian distance are noted and the bounds are chosen such that 95% of the cases are being covered. For these bounds, we plot in Figure 3 the transmitted power necessary for a BS with 6 antennas to serve 3 users positioned such that the separation angles between them is 7°. We note that our method not only requires less transmit power but also remains feasible for larger SINR values.

5. RELATED WORK AND CONTRIBUTIONS

In this paper we have proposed a new worst case robust downlink beamforming problem with imperfect covariance based CSI at the transmitter.

Previous approaches e.g., [1], [6], [12], [13] used Frobenius norms to bound the uncertainty regions around the presumed channel covariance matrix. Compared to them, our formulation is essentially different, in that we take into account the underlying Riemannian manifold structure of the space of the covariance matrices and measure the uncertainty region with a metric properly defined on this manifold. Adopting the Riemannian distance derived in [17], we find a simple upper bound which we use in our derivations. We noted the resemblance between this approximation and the bound previously used in [14]. However, the motivation in [14] is merely to simplify the positive semidefiniteness constraints on the covariance matrices, whereas in the problem formulation the uncertainty sets are still bounded based on Frobenius norm. Moreover, due to the approximations used in [14], the techniques presented there are weaker competitors to our method, as we have shown through simulations.

With respect to the recently proposed [15] where the uncertainty region is limited using trace bounds, our approach is more general in that it can be employed to a larger number of error models on the CSI, while the technique in [15] is appropriate for specific error sources, e.g., finite sampling.
6. REFERENCES


