Transmitter Design in Partially Coherent Multiple Antenna Systems

Leila Musavian, Mischa Dohler, M. Reza Nakhai, A. Hamid Aghvami  
Centre for Telecommunications Research,  
King’s College London,  
WC2R 5RL London, UK  
{leila.musavian, michael.dohler, reza.nakhai, hamid.aghvami}@kcl.ac.uk

Abstract—In this paper, we study the effect of channel estimation error on transmitter design in a multiple input multiple output (MIMO) channel. We assume that only knowledge of either the channel correlation or mean is available at the transmitter and find the transmitting strategy. Under covariance feedback, at the transmitter the channel is modelled as a matrix of zero mean circularly symmetric complex Gaussian (ZMCSCG) random variables with known covariances. We assume that only rows of channel matrix are correlated. Under mean feedback the covariance of channel is modelled as white. We determine the necessary and sufficient conditions under which a unit rank input covariance matrix can achieve the capacity lower bound. This paper is structured as follow. In section II, the system model is introduced. In Section III, we assume transmitter to be correlated and derive the conditions at which beamforming can achieve capacity lower bound. In Section IV, we solve the same problem when channel mean information is available at the transmitter. Section VI includes the numerical results and conclusions are finally drawn in section VII.

As for the notation, we will use the upper case boldface letters for matrices and lower case boldface letters for vectors; $tr(A)$ stands for the trace of a matrix $A$ and $A^H$ determines its Hermitian transpose. $A^T$ denotes the transpose of matrix $A$ and $\lambda_i^A$ is the $i^{th}$ eigenvalue of matrix $A$. The matrix (pseudo) inverse is denoted by $[\cdot]^{-1}$. $A^{1/2}$ denotes the square root of matrix $A$ and $|A|$ determines the determinant of matrix $A$. Finally $\mathcal{E}[x]$ denotes the expectation of the random variable $x$.

II. SYSTEM MODEL

We consider a wireless communication system employing $t$ transmit antennas and $r$ receive antennas. The channel is assumed to be flat with a linear model in which the received $r \times 1$ vector $y$ depends on the transmitted $t \times 1$ vector $x$ according to

$$y = Hx + n,$$  \hspace{1cm} (1)

where the entries of $r \times 1$ noise vector $n$ are independently distributed ZMCSRG noise samples with variance $\sigma_n^2$. $H$ is the $r \times t$ channel transfer matrix with ZMCSRG elements, each with unit variance. The total average transmitted power is constrained to $P$, i.e., $tr(\mathcal{E}[xx^H]) \leq P$. We restrict $x$ to have a Gaussian distribution. Note that such an input distribution in a system with imperfect CSIR is not necessarily a capacity achieving distribution [6].

A. Covariance information

Under covariance information model, We assume that the elements of each row of $H$ are correlated and follow the model described in [2], such as $H = H_w R_i^{1/2}$, where $R_i$ is the $t \times t$ transmit antenna correlation matrix. The elements of matrix $H_w$ are i.i.d.(spatially white) and ZMCSCG random variables. Note
that $R_t$ is positive semi-definite Hermitian matrix. Furthermore, the correlation matrix $R_c$ can be eigen-decomposed as $R_c = U_1 A_1 U^H_1$, where $A_1 = \text{diag}[\lambda_1^c, \ldots, \lambda_n^c]$ is a diagonal matrix with $\lambda_j^c \in (1, t)$ as the eigenvalues of $R_c$. Without loss of generality, we assume that $\lambda_1^c \geq \cdots \geq \lambda_n^c$. We further assume that the receiver has a perfect knowledge of the transmit correlation matrix and only an inaccurate information of $H_w$, which we denote as $\tilde{H}_w$. The minimum mean square error (MMSE) estimation of $H_w$ given the provided information $\tilde{H}_w$, i.e., $\tilde{H}_w = \mathcal{E}[\tilde{H}_w | H_w]$, is performed at the receiver. The MMSE estimation error matrix $E_w$ given by $E_w = H_w - \tilde{H}_w$, where $\tilde{H}_w$ and $E_w$ are uncorrelated and their entries are i.i.d. and ZMCSGC with variances $\sigma_0^2$ and $\sigma_0^2$, respectively, [6]. Hence the channel matrix can be written as [9]

$$H = \tilde{H} + E_w R_t^{1/2},$$

where $\tilde{H} = \tilde{H}_w R_t^{1/2}$. We define the total channel estimation error matrix $E = H - \tilde{H}$, which is given by $E = E_w R_t^{1/2}$.

**B. Mean information**

Under mean information model we assume the channel as

$$H = H_u + H_w,$$

where $H_u$ is the mean information(deterministic) of the channel and $H_w$ is the ZMCSGC matrix with unit variance, representing scattering. We assume the mean part of channel is known at the receiver and also an inaccurate information of $H_w$. Following the same as section II.A, we can model $H$ as

$$H = H_u + \tilde{H}_w + E_w,$$

where the total channel estimation error is $E = E_w$.

### III. TRANSMITTER DESIGN WITH COVARIANCE FEEDBACK

In this section, we use the lower bound on mutual information for correlated MIMO system such as [9, 10]

$$I(X; Y) \geq \log_2 \det \left( I_r + \frac{\tilde{H}Q\tilde{H}^H}{\sigma_n^2 + \sigma_E^2 tr(R_t Q)} \right),$$

where, $Q = \mathcal{E}[xx^H]$. We have assumed the power constraint of $tr(Q) \leq P$, at the transmitter, therefore the lower bound on capacity is given by

$$C_{lo} = \max_{Q \colon tr(Q) \leq P} \log_2 \det \left( I_r + \frac{\tilde{H}Q\tilde{H}^H}{\sigma_n^2 + \sigma_E^2 tr(R_t Q)} \right).$$

Not that with capacity lower bound $C_{lo}$, we mean the maximum of mutual information lower bound under assumption of Gaussian input signalling.

We assume that the channel covariance information is available at the transmitter. It has been proved that the transmission scheme in MIMO channels with these assumptions will end up to be an eigen-beamformer [9, 10]. In the case when CSI is imperfect, we generalize the conditions at which beamforming achieves the capacity lower bound.

**Theorem 1**: For the case of uncorrelated receive antennas the maximum on mutual information lower bound can be achieved by beamforming for given eigenvalues of transmit correlation matrix, if and only if the following inequality is fulfilled:

$$\frac{\sigma_n^2}{P\lambda^2} > \frac{r(1 - \sigma_E^2)}{1 - \left( \frac{\sigma^2_E + \sigma_n^2}{\sigma_n^2} \right) \left( 1 - r \right)},$$

where $\Gamma(a, x) \triangleq \int_x^\infty t^{-a} e^{-t} dt$ is the upper incomplete Gamma function.

**Proof**: When the beamforming is the transmitting strategy, then the power allocated to the eigenvectors are $[P, 0, \cdots, 0]$. This means that the whole power is allocated to the eigenvector associated to the largest eigenvalue of the channel. Now, let us assume that the amount of power allocated to the eigenvector with the largest eigenvalue is $P - \rho$ and the remaining amount of power $\rho$ is distributed among the other eigenvectors, such that $\lambda_i^Q = \alpha_i \rho$, $i \in (2, 3, \cdots, t)$, where $\sum_{i=2}^t \alpha_i = 1$ and $\alpha_i \geq 0$. Then, to find the condition for beamforming to achieve $C_{lo}$, we have to find the conditions that $C_{lo}(\rho) \leq C_{lo}(0)$ for all $0 \leq \rho \leq P$. Defining $\tilde{H}_w, \tilde{H}_w^t = Z_i$, where $\tilde{H}_w$ is the $i^{th}$ column of the matrix $\tilde{H}_w$, we derive the capacity lower bound under these conditions as follows [10]

$$C_{lo}(\rho) = \log_2 \left| I_r + \frac{(P - \rho)\lambda_i^1 Z_i + \rho \sum_{i=2}^t \alpha_i \lambda_i^Z_i}{\sigma_n^2 + (P - \rho)\sigma_E^2 \lambda_i^1 + \rho \sigma_E^2 \sum_{i=2}^t \alpha_i \lambda_i^1} \right|$$

$$= \log_2 \left| I_r + \frac{(P - \rho)\lambda_i^1 Z_i + \rho \sum_{i=2}^t \alpha_i \lambda_i^Z_i}{\sigma_n^2 + P\sigma_E^2 \lambda_i^1} \left[ 1 + \frac{\rho}{\sigma_n^2 + P\sigma_E^2 \lambda_i^1} \right] \right|.$$

To simplify the subsequent inequality, we define $b \triangleq (\sigma_n^2 + P\sigma_E^2 \lambda_i^1)$ and $\frac{\rho}{\sigma_n^2 + P\sigma_E^2 \lambda_i^1} \triangleq (\sigma_n^2 + \sigma_E^2 \lambda_i^1)$. Using the equation $(1 + x)^{-1} = (1 - x + x^2 - x^3 + \cdots)$ [11], we can approximate $(1 + x)^{-1} \approx (1 - x)$ for $|x| < 1$, we approximate the capacity lower bound as

$$C_{lo}(\rho) \approx \log_2 \left| I_r + \frac{(P - \rho)\lambda_i^1 Z_i + \rho \sum_{i=2}^t \alpha_i \lambda_i^Z_i}{\sigma_n^2 + \frac{P\lambda_i^1}{a} Z_i} \right|$$

$$= \log_2 C_{lo} + \rho D + \rho^2 F,$$

where

$$C = I_r + \frac{P\lambda_i^1}{a},$$

$$D = \frac{1}{a} \left( \sum \alpha_i \lambda_i^1 Z_i - (1 + Pb) \lambda_i^1 Z_i \right),$$

$$F = -\frac{b}{a} \left( \sum \alpha_i \lambda_i^1 Z_i - \lambda_i^1 Z_i \right).$$

To prove that $C_{lo}(\rho)$ is maximum at $\rho = 0$ for $0 \leq \rho \leq P$, we first need to find the region at which $\frac{\partial C_{lo}(\rho)}{\partial \rho} \bigg|_{\rho=0} \leq 0$.

Using the equality [10]

$$\frac{\partial \log_2 \det (X)}{\partial z} = tr \left( X^{-1} \frac{\partial X}{\partial z} \right).$$

0-7803-8939-5/05/$20.00 (C) 2005 IEEE
We find the differentiation of $C_{lo}(\rho)$ as follows

\[
\frac{\partial C_{lo}(\rho)}{\partial \rho} \Bigg|_{\rho=0} = \mathcal{E} \left[ tr \left( (C + \rho D + \rho^2 F)^{-1} \right) \right] \] (13)

\[
= \mathcal{E} \left[ tr(\mathbb{C}) \right] \]

\[
= \frac{1}{a} \mathcal{E} \left[ tr \left( \left( I_r + \frac{P \lambda_i}{a} Z_i \right)^{-1} \alpha_i \lambda_i Z_i \right) \right] - \frac{1}{a} \mathcal{E} \left[ tr \left( \left( I_r + \frac{P \lambda_i}{a} Z_i \right)^{-1} \left( 1 + P b \right) \lambda_i Z_i \right) \right] \] (15)

Following the same steps as [4], we can simplify (15) as

\[
\frac{\partial C_{lo}(\rho)}{\partial \rho} \Bigg|_{\rho=0} = \sum_{i=2}^{t} - \frac{\sigma^2}{P \alpha} \left( r(1 - \sigma^2) - 1 + \mathcal{E} \left[ \left( 1 + P \frac{tr(Z_i)^2}{a} \right)^{-1} \right] \right) \] (17)

It is clear from (17) that the best distribution for $\alpha$ to maximize $\frac{\partial C_{lo}(\rho)}{\partial \rho} \bigg|_{\rho=0}$ is when $\alpha_2 = 1$ and $\alpha_i = 0$ for $i \in (3, t)$, or equally when the power $\rho$ is totally transmitted along the eigenvector associated to the second largest eigenvalue. Using the closed form evaluation of the expectation [4]

\[
\mathcal{E} \left[ \frac{1}{1 + \frac{P \lambda_i^2}{a} tr(Z_i)} \right] = \left( \frac{1}{1 - \sigma^2} \right) \frac{1}{P \lambda_i^2} \Gamma \left( 1 - r, \frac{a}{1 - \sigma^2} \right) \] (18)

in (17), we finally proved the necessity of the Theorem 1.

In order to prove that this condition is also sufficient for beamforming to achieve capacity lower bound, we use the equation $\frac{\partial^2 C_{lo}(\rho)}{\partial \rho^2} = \mathcal{E} \left[ tr(\mathbb{N}) \right]$, where $\mathbb{N} = (C + \rho D + \rho^2 F)^{-1} (D + 2 \rho F)$ [10] and show that $\frac{\partial^2 C_{lo}(\rho)}{\partial \rho^2}$ is non-positive for all values of $\rho$ irrespective of the channel conditions. Following the same simplification method in [4], since $(C + \rho D + \rho^2 F)^{-1}$ is a Hermitian matrix, then $\frac{\partial^2 C_{lo}(\rho)}{\partial \rho^2}$ can be obtained as $\frac{\partial^2 C_{lo}(\rho)}{\partial \rho^2} = - \mathcal{E} \left[ tr(M^2) \right]$, where $M = (C + \rho D + \rho^2 F)^{-1/2} (C + 2 \rho D) (C + \rho D + \rho^2 F)^{-1/2}$ is a positive semi-definite matrix. Therefore, $\frac{\partial^2 C_{lo}(\rho)}{\partial \rho^2} \leq 0$ for all $\rho$, $0 < \rho < P$. Thus the necessary condition (7) is also a sufficient condition for beamforming to be the transmit strategy.

In the case that there is perfect channel estimation at the receiver, i.e. $(\sigma^2 E = 0)$, we get the same conditions as were already derived in [4] and [5].

IV. TRANSMITTER DESIGN WITH MEAN FEEDBACK

In this section, in order to find the lower bound on mutual information, we use equations (4) and (5), when the transmit covariance matrix is identity matrix and $\mathbf{H} = \mathbf{H}_\mu + \hat{\mathbf{H}}_w$, therefore

\[
\mathbf{I}(\mathbf{X}; \mathbf{Y}) \geq \log_2 \left[ \frac{1}{1 - \sigma^2} \left( \mathbf{I}_r + \frac{\mathbf{H}_\mu + \hat{\mathbf{H}}_w}{\sigma^2_n + P \sigma^2_E} \right) \right], \quad (19)
\]

we can see that in systems with uncorrelated transmitter antennas, the worst effect of channel estimation error is the same as an additive white Gaussian noise (AWGN). Using (19), the lower bound on capacity under the power constraint can be written as

\[
C_{lo} = \max_{\mu} \mathbb{E} \left[ tr(\mathbb{Q}) \right] \log_2 \left[ \frac{1}{1 - \sigma^2} \left( \mathbf{I}_r + \frac{\mathbf{H}_\mu + \hat{\mathbf{H}}_w}{\sigma^2_n + P \sigma^2_E} \right) \right]. \quad (20)
\]

Comparing to the optimization problem solved in [4], we can see that the optimized transmit strategy under mean feedback used in coherent system is also the optimum based on capacity lower bound in the system with uncertainty in CSIR. Thus the input covariance matrix that solves (20) is $\mathbf{U}_\mu = U_{\mu}^o$, where the first column of matrix $\mathbf{U}_\mu^o$ is $\mathbf{H}_\mu^o \mathbf{U}_\mu^o [1]$, where $\mathbf{H}_\mu^o [1]$ is the first row of matrix $\mathbf{H}_\mu$. The other columns of the unitary matrix $\mathbf{U}_\mu^o$ are arbitrary orthonormal vectors.

To find the conditions for the beamforming being the transmitting solution, we use decomposition of $\mathbb{Q}$ in (20) such as

\[
C_{lo} = \mathcal{A}_Q : \mathbb{E} \left[ tr(\mathbb{Q}) \right] \leq \log_2 \left[ \frac{1}{1 - \sigma^2} \left( \mathbf{I}_r + \frac{\mathbf{Z}Q\mathbf{Z}^H}{\sigma^2_n + P \sigma^2_E} \right) \right]. \quad (21)
\]

where $\mathbf{Z} = \frac{1}{\sqrt{1 - \sigma^2}} (\mathbf{H}_\mu \mathbf{U}_\mu^o + \hat{\mathbf{H}}_w)$ we proceed the same as last section to get

\[
C_{lo}(\rho) = \log_2 \left[ \mathbf{I}_r + (P - \rho) \frac{(1 - \sigma^2)}{\sigma^2_n + P \sigma^2_E} \right] + \ldots \right] \quad (22)
\]

We need to find the condition under which $\frac{\partial C_{lo}(\rho)}{\partial \rho} \bigg|_{\rho=0} \leq 0$.

Following thorough the steps (16) to (20), the necessary condition of beamforming to achieve the capacity lower bound can be found as

\[
\mathcal{E} \left[ \left( 1 - \frac{1}{\sigma^2_n + P \sigma^2_E} \right) \left( 1 - \frac{1}{\sigma^2_n + P \sigma^2_E} \right) \right] \leq \frac{1 + \frac{P(1 - \sigma^2)}{\sigma^2_n + P \sigma^2_E}}{1 + \frac{P}{\sigma^2_n + P \sigma^2_E}}, \quad (23)
\]

where $w$ is a non-central chi-squared distributed random variable with con-centrality parameter $\delta = \frac{\sigma^2}{\frac{\sigma^2}{\sigma^2 + P}}$ and $2r$ degrees of freedom.

To proof that this condition is also sufficient, the same approach as covariance feedback can be applied.
V. NUMERICAL RESULTS

For our simulations, we consider various number of transmit and receive antennas in a Rayleigh fading channel. We assume linear antenna arrays at the transmitter, where its elements are respectively spaced by $d_t$. The receiver is assumed to be uniformly surrounded by clutter, while the transmitter is unobstructed by local scatterer. This is appropriate to the “one-ring” model used in [12]. This model uses the distance $D$ between the transmit and the receive antennas and the radius $R$ of the scatter ring as parameters. Based on this model and the results from [12], the entries of the transmit and receive correlation matrices are given by

$$R_{tx}(k,l) = J_0\left(2\pi|k-l|\frac{d_t}{\lambda_c}\right)$$

where $\lambda_c$ is the carrier wavelength, $J_0$ is zeroth-order Bessel function of the first kind. In the mean information case, we assume that the spacing between the antenna elements at receiver is large enough to achieve a de-correlated receiver side.

In Fig. 1, we plot the lower bounds of mutual information of a MIMO channel, when only transmitter side is correlated and the input power covariance matrix is a scaled identity matrix, i.e. $Q = \frac{\sigma_e^2}{\lambda_c^2}I_t$. In this plot $\sigma_e^2$ is assumed to be 0.1 and both bounds are drawn versus SNR. This figure shows that in the presence of CSI error the mutual information curves saturate at high SNRs. It is happened because we have assumed that the channel estimation quality is not improving when SNR increases.

In Fig. 2, we draw the conditions in Theorem 1 for a different number of receive antennas and different error variances. Note that the region below each line presents the region which beamforming can achieve the maximum on mutual information lower bound, i.e. $C_{lo}$. The chance for beamforming to achieve $C_{lo}$ is bigger when this region is bigger. We can see from Fig. 2 that in the presence of CSI error, the chance for beamforming to achieve the maximum on mutual information lower bound increases, or equivalently, a system with less correlated transmitter can provide the condition in Theorem 1 when CSI is imperfect at the receiver.

The more CSI error occurs at the receiver, the less correlation is needed at the transmitter for a unit rank input matrix to achieve $C_{lo}$.

In Fig. 3, we plot the difference between regions where beamforming achieves $C_{lo}$ when there is no error and when there is an error equal to $\sigma_e^2$ to illustrate the effect of channel estimation error on the beamforming region. Here, a system with uncorrelated receiver is assumed and we observe that the probability for beamforming to achieve the maximum on mutual information lower bound increases when $\sigma_e^2$ increases.

Finally, in Fig. 4, the system is assumed to have only one receive antennas. Here, we assumed that only a perfect information from the channel mean $(\mu_k)$ is known to the transmitter. Therefore, $h \sim N(\mu_k,(1-\sigma_e^2)I_t)$. We find the value of $\mu^2_k$ for which beamforming maximizes the mutual information lower bound for different $\sigma_e^2$. We can see from the figure that when channel estimation error increases, the probability that beamforming can achieve the capacity lower bound increases.

VI. CONCLUSIONS

In this paper, we have studied the effect of uncertainty on the channel knowledge at the receiver in a single user MIMO system. We used lower bounds on mutual information when the knowledge on the error variance was available. We assume that either the channel correlation or mean information is perfectly known to the receiver and solve the transmitter design problem when this information is available at the transmitter perfectly.

In this case, we proved that the optimal transmit strategy when there is an ideal channel estimation at the receiver, is also optimal in partially coherent systems based on capacity lower bound. Furthermore, the necessary and sufficient conditions for input covariance matrix to have unity rank, i.e. eigen-beamforming, is generalized and new conditions are derived, taking to account the CSI estimation error at the receiver. It was shown that, in general, a higher error in the channel increases the probability of beamforming to achieve the capacity lower bound.

Fig. 3. Difference between the necessary and sufficient condition for beamforming to achieve the maximum on mutual information lower bound when there is perfect and imperfect CSI in a MIMO system uncorrelated receive antennas.

Fig. 4. Necessary and sufficient condition for beamforming to maximize the lower bound on mutual information with mean feedback.

REFERENCES