

A PROBABILISTIC CONSTRAINT APPROACH FOR ROBUST TRANSMIT BEAMFORMING WITH IMPERFECT CHANNEL INFORMATION

Pei-Jung Chung¹, Huiqin Du¹, Jacek Gondzio²

¹School of Engineering and Electronics, ²School of Mathematics

The University of Edinburgh, UK

P.Chung@ed.ac.uk, H.Du@ed.ac.uk, J.Gondzio@ed.ac.uk

ABSTRACT

Transmit beamforming is a powerful technique for enhancing performance of wireless communication systems. Most existing transmit beamforming techniques require perfect channel state information at the transmitter (CSIT), which is typically not available in practice. In such situations, the design should take errors in CSIT into account to avoid performance degradation. Among two popular robust designs, the stochastic approach exploits channel statistics and optimizes the average system performance. The maximin approach considers errors as deterministic and optimizes the worst-case performance. The latter usually leads to conservative results as the extreme (but rare) conditions may occur at a very low probability. In this work, we propose a more flexible approach that maximizes the average signal-to-noise ratio (SNR) and takes the extreme conditions into account proportionally. Simulation results show that the proposed beamformer offers higher robustness against channel estimation errors than several popular transmit beamformers.

1. INTRODUCTION

Multi-antenna diversity is well motivated in wireless communication systems because it offers significant advantages over single antenna [1]. Perfect or partial knowledge of the channel state information at transmitter (CSIT) can provide further performance improvement.

However, in practical wireless systems, accurate channel estimates are not available due to errors induced by imperfect channel feedback, estimation/quantization errors or outdated channels. It is well known that the performance of several nonrobust designs for multi-antenna diversity degrades rapidly with increasing error levels. This has motivated many works that take imperfect channel information into account.

Existing robust transmit beamforming (or precoder) designs can be categorized into the stochastic and the maximin approaches. The stochastic approach [2] [3] exploits channel statistics such as mean or covariance and optimizes the average system performance. On the other hand, the maximin approach considers channel estimation errors as deterministic and optimizes the worst-case performance [4] [5]. While the stochastic approach focuses on the average performance without paying attention to the extreme error level, the worst-case approach is overall too conservative as the worst operational condition is rare.

To overcome this problem, we proposed a more flexible design based on probabilistic constraint using channel covariance in [6]. In this work, we apply this approach to transmit beamforming design under consideration of

imperfect channel estimates. Note that a similar strategy was introduced into the design of adaptive beamformer at the receiver side in [7].

Our approach maximizes the average Signal-to-Noise Ratio (SNR) and ensures robustness against the CSIT error by keeping the probability of the worst-case performance at a very low level. Under the assumption that the CSIT error is complex Gaussian distributed, this stochastic optimization problem is further simplified to an equivalent deterministic form which can be efficiently solved by modern convex optimization algorithms [8]. Simulation results show that the proposed approach provides the best performance and highest robustness among several popular transmit beamformers.

In the following section, we give a brief description of the system model. The proposed approach is formulated as a stochastic optimization problem in Section 3. Section 4 is devoted to transformation of the probabilistic constraint to a deterministic, convex constraint. Simulation results are presented in Section 5. Finally, Section 6 concludes this paper.

2. SYSTEM MODEL

Consider a single-user wireless communication system with N_t transmit antennas and N_r receive antennas. The encoded signal $\mathbf{s} \in \mathbb{C}^{P \times 1}$ is spread by the precoding matrix $\mathbf{C} \in \mathbb{C}^{N_t \times P}$ and then transmitted through a flat fading channel. The received signal \mathbf{y} in the presence of additive white Gaussian noise \mathbf{w} is given by

$$\mathbf{y} = \mathbf{H}\mathbf{C}\mathbf{s} + \mathbf{w}. \quad (1)$$

The (i, j) element of the channel matrix $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_{N_t}] \in \mathbb{C}^{N_r \times N_t}$ represents the response between the i th receive antenna and the j th transmit antenna. Assuming perfect channel knowledge at the receiver, the average signal-to-noise-ratio (SNR) obtained from maximum ratio combining (MRC) is given by

$$\text{SNR} = \frac{E_s}{N_0} \text{tr}\{\mathbf{C}^H \mathbf{H}^H \mathbf{H} \mathbf{C}\}, \quad (2)$$

where $E_s = \mathbb{E}[\|\mathbf{s}\|^2]$ is the average energy of the signal and N_0 is the noise power.

When perfect channel knowledge is available at transmitter, maximization of the average SNR leads to the conventional one directional beamforming which allocates all power on the strongest eigen-mode of the channel correlation matrix $\mathbf{H}^H \mathbf{H}$. In practice, one has only access to an imperfect estimate for the channel matrix $\hat{\mathbf{H}} \in \mathbb{C}^{N_r \times N_t}$, which is related to \mathbf{H} as follows:

$$\mathbf{H} = \hat{\mathbf{H}} + \mathbf{E}, \quad (3)$$

where the error matrix $\mathbf{E} \in \mathbb{C}^{N_r \times N_t}$ consists of i.i.d. complex normally distributed entries with variance σ_e^2 . The goal of this work is to design a transmit beamformer \mathbf{C} that maximizes SNR under consideration of inaccuracy in channel estimates.

3. ROBUST DESIGN BASED ON PROBABILISTIC CONSTRAINED OPTIMIZATION

To tackle performance degradation caused by imperfect channel estimates, we consider a probabilistic constraint approach. The proposed algorithm maximizes the average SNR while keeping the probability for SNR being below a pre-specified threshold γ_{th} low. It has the advantage of achieving optimal overall performance while providing quality control for the worst case. In contrast to the mini-max approach that focuses on the worst-case performance, the probability constraint takes the errors into account proportionally. On the other hand, the worst case scenario ignored by the stochastic approach is considered in our approach.

Assuming the error model (3), the average SNR (2) becomes a function of the channel estimate $\hat{\mathbf{H}}$ and the random error \mathbf{E}

$$f(\hat{\mathbf{H}}, \mathbf{E}) = \frac{E_s}{N_0} \text{tr}\{\mathbf{C}^H (\hat{\mathbf{H}} + \mathbf{E})^H (\hat{\mathbf{H}} + \mathbf{E}) \mathbf{C}\}. \quad (4)$$

To simplify the expression (4), we consider the eigen-decomposition of $\mathbf{C}\mathbf{C}^H = \mathbf{U}_c \mathbf{D}_c \mathbf{U}_c^H$ and $\hat{\mathbf{H}}^H \hat{\mathbf{H}} = \mathbf{U}_h \mathbf{D}_h \mathbf{U}_h^H$. The diagonal matrix $\mathbf{D}_c = \text{diag}(d_1, d_2, \dots, d_{N_t})$ where $d_1 \geq \dots \geq d_{N_t} \geq 0$ are eigenvalues of $\mathbf{C}\mathbf{C}^H$. The corresponding eigenvectors are summarized in the unitary matrix \mathbf{U}_c . The matrices $\mathbf{D}_h = \text{diag}(D_1, \dots, D_{N_t})$ and \mathbf{U}_h are similarly defined.

3.1 Objective function

Given the channel estimate $\hat{\mathbf{H}}$, we obtain the objective function by taking the expectation of $f(\hat{\mathbf{H}}, \mathbf{E})$ with respect to the random error \mathbf{E}

$$\mathbb{E} [f(\hat{\mathbf{H}}, \mathbf{E})] = \frac{E_s}{N_0} \text{tr}\{\mathbf{U}_c \mathbf{D}_c \mathbf{U}_c^H (\mathbf{U}_h \mathbf{D}_h \mathbf{U}_h^H + \sigma_e^2 N_r \mathbf{I}_{N_t})\}. \quad (5)$$

It is well established in the literature [4] that a function with a structure similar to (5) can be maximized over \mathbf{U}_c and \mathbf{D}_c separately. Inserting the optimal solution for \mathbf{U}_c so that $\mathbf{U}_c^H \mathbf{U}_h = \mathbf{I}$, we obtain the following objective function

$$\bar{f}(\mathbf{D}_c) = \frac{E_s}{N_0} \text{tr}\{\mathbf{D}_c (\mathbf{D}_h + \sigma_e^2 N_r \mathbf{I}_{N_t})\}. \quad (6)$$

Note that $\bar{f}(\mathbf{D}_c)$ depends on $\mathbf{C}\mathbf{C}^H$ only through its eigenvalues. Hence, the design of the beamforming matrix becomes a power allocation problem.

3.2 Probabilistic constraint

To mitigate the impact of large errors, we guarantee the system performance by keeping the probability that SNR becomes smaller than an acceptable level γ_{th} to be low. More precisely, given an acceptable SNR level γ_{th} and the outage probability p_{out} , $f(\hat{\mathbf{H}}, \mathbf{E})$ satisfies the following probabilistic constraint

$$\Pr\{f(\hat{\mathbf{H}}, \mathbf{E}) \leq \gamma_{th}\} \leq p_{out}, \quad (7)$$

where $\Pr\{A\}$ denotes the probability of the event A .

As shown in (7), the distribution of $f(\hat{\mathbf{H}}, \mathbf{E})$ is crucial to the implementation of our algorithm. Applying the eigen-decomposition of $\mathbf{C}\mathbf{C}^H = \mathbf{U}_c \mathbf{D}_c \mathbf{U}_c^H$ and permutation property of the trace operation, (7) can be simplified to a mixture of independent noncentral $\chi_{n_i}^2(\delta_i)$ -distributed random variables Z_i , $i = 1, \dots, N_t$

$$f(\hat{\mathbf{H}}, \mathbf{E}) = \frac{E_s}{N_0} \sum_{i=1}^{N_t} d_i \sigma_e^2 Z_i. \quad (8)$$

The noncentrality parameter is $\delta_i = \frac{1}{\sigma_e^2} \tilde{\mathbf{h}}_i^H \tilde{\mathbf{h}}_i$ and the degree of freedom is $n_i = 2N_r$. The vector $\tilde{\mathbf{h}}_i \in \mathbb{C}^{N_r \times 1}$ represents the i th column of the matrix $\tilde{\mathbf{H}} = \hat{\mathbf{H}} \mathbf{U}_c$.

3.3 Probabilistic constrained optimization

Having derived the average SNR (5) and the compact expression (8) for $f(\hat{\mathbf{H}}, \mathbf{E})$, our design can be formulated as the following constrained optimization problem:

$$\max_{\mathbf{D}_c} \text{tr}\{\mathbf{D}_c (\mathbf{D}_h + \sigma_e^2 N_r \mathbf{I}_{N_t})\},$$

subject to

$$\Pr\left\{\sum_{i=1}^{N_t} d_i Z_i \leq \bar{\gamma}\right\} \leq p_{out}, \quad (9)$$

$$\text{tr}\{\mathbf{D}_c\} \leq 1, \quad (10)$$

$$d_i \geq 0, \quad i = 1, \dots, N_t \quad (11)$$

where $\bar{\gamma} = \gamma_{th} (\frac{E_s}{N_0} \sigma_e^2)^{-1}$ and (10) is a convex constraint derived from the power constraint $\text{tr}\{\mathbf{C}\mathbf{C}^H\} \leq 1$.

4. REFORMULATION OF PROBABILISTIC CONSTRAINT

The major challenge in our approach is to convert the probabilistic constraint (9) into a deterministic one so that the solution can be efficiently computed by standard tools of mathematical programming. When the chance constraint involves linear combination of normally distributed random variables, it can be reformulated as a convex constraint [9]. However, (9) involves a mixture of noncentral χ^2 -distributions. The following result shows that the probabilistic constraint (9) can be replaced by a deterministic convex constraint.

Proposition The probabilistic constraint (9) can be replaced by the following convex constraint

$$\prod_{i=1}^{N_t} \left(\frac{1}{d_i} \left[\frac{\bar{\gamma}/2}{1 + \delta_i/n_i} \right] \right)^{n_i/2} \leq p_{out}, \quad (12)$$

where $\bar{\gamma} = \gamma_{th} (\frac{E_s}{N_0} \sigma_e^2)^{-1}$, $\delta_i = \frac{1}{\sigma_e^2} \tilde{\mathbf{h}}_i^H \tilde{\mathbf{h}}_i$ and $n_i = 2N_r$. If (12) holds, then (9) holds.

Proof: To decouple the design parameter d_i , we exploit the independence of Z_i , $i = 1, \dots, N_t$. Define the event

$$\mathcal{A}_i = \{d_i Z_i \leq \bar{\gamma}\} \quad (13)$$

and

$$\mathcal{A} = \left\{ \sum_{i=1}^{N_t} d_i Z_i \leq \bar{\gamma} \right\}. \quad (14)$$

By definition, \mathcal{A} is a subset of the intersection of \mathcal{A}_i , $i = 1, \dots, N_t$,

$$\mathcal{A} \subset \mathcal{B} = \mathcal{A}_1 \cap \mathcal{A}_2 \cap \dots \cap \mathcal{A}_{N_t} \quad (15)$$

which leads to the following inequality

$$\Pr\{\mathcal{A}\} \leq \Pr\{\mathcal{B}\} = \prod_{i=1}^{N_t} \Pr\{\mathcal{A}_i\}. \quad (16)$$

The above expression has the advantage that the event \mathcal{A}_i depends only on the noncentral $\chi_{n_i}^2(\delta_i)$ -distribution.

According to [10], the distribution of noncentral χ^2 -distribution can be approximated by a central χ^2 -distribution. Application of this result leads to the following approximation

$$\Pr\{\chi_{n_i}^2(\delta_i) \leq \frac{\bar{\gamma}}{d_i}\} \approx \Pr\{\chi_{n_i}^2 \leq \frac{\bar{\gamma}/d_i}{1 + \delta_i/n_i}\}. \quad (17)$$

To transform (17) to a deterministic form, we apply the sharp upper bound on the integral $\frac{1}{\Gamma(1+1/u)} \int_x^\infty e^{-t} t^u dt$ derived in [11] to obtain the following inequality

$$\Pr\{\chi_{n_i}^2 \leq \frac{\bar{\gamma}/d_i}{1 + \delta_i/n_i}\} < \left(1 - \exp\left(-\frac{\bar{\gamma}/d_i}{2(1 + \delta_i/n_i)}\right) \right)^{\frac{n_i}{2}}. \quad (18)$$

Due to limited space, details about the derivation of (18) will be given in a future publication.

To achieve a convex constraint, we apply the following inequality to the right hand side of (18)

$$\left(1 - \exp\left(-\frac{\bar{\gamma}/d_i}{2(1 + \delta_i/n_i)}\right) \right)^{\frac{n_i}{2}} \leq \left(\frac{\bar{\gamma}/d_i}{2(1 + \delta_i/n_i)} \right)^{\frac{n_i}{2}}. \quad (19)$$

Eq (19) follows immediately from the inequality $(1 - e^{-u})^x \leq u^x$ for $u = \frac{\bar{\gamma}/d_i}{2(1 + \delta_i/n_i)} \geq 0$ and $x = \frac{n_i}{2} > 0$. Moreover, it is easy to verify that for positive $\frac{n_i}{2}$ and $\frac{\bar{\gamma}/d_i}{2(1 + \delta_i/n_i)}$, (12) is a convex set in d_i s.

Combing the inequalities (16),(17),(18) and (19), we conclude that the probabilistic constraint (9) is satisfied by the convex constraint (12). \square

Replacing the probabilistic constraint (9) with the deterministic constraint (12), the original problem is transformed to the following convex optimization problem

$$\max_{\mathbf{D}_c} \text{tr}\{\mathbf{D}_c(\mathbf{D}_h + \sigma_e^2 N_r \mathbf{I}_{N_t})\},$$

subject to

$$\prod_{i=1}^{N_t} \left(\frac{1}{d_i} \left[\frac{\bar{\gamma}/2}{1 + \delta_i/n_i} \right] \right)^{n_i/2} \leq p_{\text{out}},$$

$$\text{tr}\{\mathbf{D}_c\} \leq 1,$$

$$d_i \geq 0, \quad i = 1, \dots, N_t,$$

that can be efficiently solved by standard tools of mathematical programming.

5. SIMULATION

In this section, we present simulation results to demonstrate robustness of the proposed beamformer in various scenarios. Here a single-user MIMO system with $N_t = 4$ transmit antennas and $N_r = 3$ receive antennas is considered. We also compare the proposed beamformer with existing techniques, such as the conventional one-directional beamformer, two-directional, equal-power loading beamformer [1] and the robust minimax beamformer [4]. We choose [4] for comparison because it uses the same type of channel information. The outage probability p_{out} is 10% and the normalized SNR threshold $\bar{\gamma}$ is 0.9 in all experiments.

In the first experiment, the proposed approach is applied to a well conditioned channel $\mathbf{D}_h = \text{diag}(0.8064, 0.1901, 0.0035, 0)$ with the first eigenvalue much larger than the remaining eigenvalues. The error variance σ_e^2 varies from 0 to 1. The SNR averaged over 10^4 Monte Carlo trials is plotted in Fig 1. With increasing error levels, the performance of all beamforming techniques degrade. For σ_e^2 between 0 and 0.4, the proposed approach, the maximin approach [4] and the one directional beamformer perform similarly. For $\sigma_e^2 > 0.4$, our approach has a much slower decline in SNR than other beamformers. The equal power loading beamformer has a significantly lower SNR than other three beamformers because channel information is not fully incorporated in its design.

In the second experiment, we consider the channel condition $\mathbf{D}_h = \text{diag}(0.4676, 0.4104, 0.1220, 0)$ with two closely spread eigenvalues. This indicates a larger correlation between antennas. Fig 2 shows that the probabilistic constraint beamformer still outperforms other three beamformers. For large error region $0.5 < \sigma_e^2 < 1$, the performance of all beamformers degrade more rapidly than in the previous experiment. However, our approach shows least sensitivity to channel errors.

In Fig 2, we also observe that when σ_e^2 increases from 0 to 0.8, the decrease in SNR associated with our approach is $\Delta SNR = -0.38$ dB and other beamformers lead to a decrease of more than $\Delta SNR = -1.83$ dB. This is 4.8 times as high as that caused by the probabilistic constraint approach. On the other hand, given a target SNR level, for example, -0.8 dB, the probabilistic constraint beamformer has the largest error tolerance range, $\sigma_e^2 \in [0, 0.78]$, while the worst case design achieves the desired performance only for $\sigma_e^2 \in [0, 0.63]$.

To summarize, the probabilistic constraint approach outperforms the worst case design and other classical beamformers over the entire error range. The gain in SNR is most significant at high error levels. In other words, our approach has the broadest tolerance range for channel

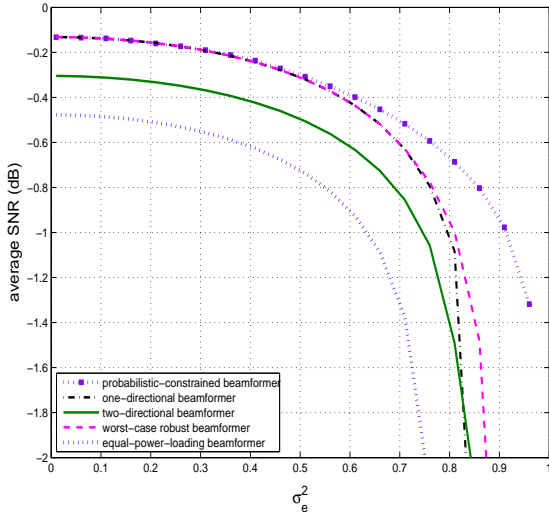


Figure 1: Average SNR versus σ_e^2 . $p_{\text{out}} = 10\%$, $\bar{\gamma} = 0.9$, $\mathbf{D}_h = \text{diag}(0.8064, 0.1901, 0.0035, 0)$.

errors. Since the original stochastic optimization problem has been transformed to a convex optimization problem, the computational complexity is similar to the worst-case approach such as [4].

6. CONCLUSION

We proposed a novel transmit beamforming design that maximizes the average SNR performance and also guarantees robustness against channel estimation errors. Our approach was formulated as a probabilistic constrained optimization problem. Under the assumption that the channel estimation error is complex Gaussian distributed, the underlying problem was transformed into a convex optimization problem which can be efficiently solved by modern software packages. The resulting computational cost is similar to many state-of-the-art robust transmit beamformers. Simulation results show that the proposed beamformer achieves higher robustness than the maximin approach and leads to a much broader tolerance range for channel estimation errors. It provides a promising alternative to existing robust transmit beamforming techniques.

REFERENCES

- [1] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, Cambridge University Press, first edition, 2003.
- [2] S. Zhou and G.B. Giannakis, "Optimal transmitter eigen-beamforming and space-time blocking based on channel mean feedback," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2599–2613, July 2002.
- [3] S. Zhou and G.B. Giannakis, "Optimal transmitter eigen-beamforming and space-time blocking based on channel covariance," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1673–1690, July 2003.
- [4] A. Pascual-Iserte, D. P. Palomar, A. I. Perez-Neira, and M. A. Lagunas, "A robust maximin approach for MIMO

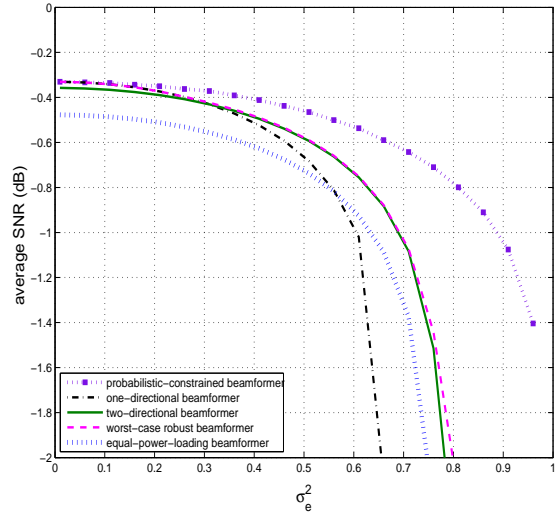


Figure 2: Average SNR versus σ_e^2 . $p_{\text{out}} = 10\%$, $\bar{\gamma} = 0.9$, $\mathbf{D}_h = \text{diag}(0.4676, 0.4104, 0.122, 0)$.

communications with imperfect channel state information based on convex optimization," *IEEE Transactions on Signal Processing*, vol. 54, no. 1, pp. 346–360, January 2006.

- [5] A. Abdel-Samad, T. N. Davidson, and A. B. Gershman, "Robust transmit eigen beamforming based on imperfect channel state information," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1596–1609, May 2006.
- [6] H. Q. Du, P.-J. Chung, J. Gondzio, and B. Mulgrew, "Robust transmit beamforming based on probabilistic constraint," in *Proc. EUSIPCO*, Lausanne, Switzerland, August 2008.
- [7] S. A. Vorobyov, Y. Rong, and A. B. Gershman, "Robust adaptive beamforming using probabilistic-constrained optimization," in *Proc. IEEE Statistical Signal Processing Workshop*, 2005.
- [8] M. Grant, S. Boyd, and Y. Ye, "CVX User's Guide for cvx version 1.1," November 2007.
- [9] P. Kall and S. W. Wallace, *Stochastic Programming*, John Wiley & Sons, 1994.
- [10] D. R. Cox and N. Reid, "Approximation to noncentral distribution," *The Canadian Journal of Statistics*, vol. 15, no. 2, pp. 105–114, 1987.
- [11] H. Alzer, "On some inequalities for the incomplete gamma function," *Mathematics of Computation*, vol. 66, no. 218, pp. 771–778, April 1997.