Covariance-Based Signaling and Feedback Data Parameterization for the Time-Varying MIMO Broadcast Channel

Adam L. Anderson[1], James R. Zeidler[1], Michael A. Jensen[2]

[1]Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407, USA, Email: adam_anderson@aj-design.com, zeidler@ece.ucsd.edu

[2]Department of Electrical and Computer Engineering, Brigham Young University, 459 CB, Provo, UT 84602, USA, Email: jensen@ee.byu.edu

Abstract

Linear precoding (beamforming) techniques that maximize the sum rate in the multi-antenna broadcast channel often suffer severe performance degradation when the channel state information at the transmitter is erroneous or outdated. This paper presents a linear precoding algorithm that uses the channel statistics rather than channel state to provide more stable performance in time-varying environments. This stable beamforming method, however, requires that the transmitter know the full spatial correlation matrix, significantly increasing the amount of data communicated over the feedback channel. Parameterization of the spatial correlation using popular channel models is therefore considered to reduce the size of the feedback data stream.

1. Introduction

Traditional signaling strategies for the multi-user multiple-input multiple-output (MIMO) wireless broadcast channel such as the optimal nonlinear dirty-paper coding (DPC) [1] or the optimal linear regularized channel inversion (RCI) beamformer [2, 3] require accurate channel state information (CSI) at the transmitter (CSIT) and receiver (CSIR). It has previously been shown that both DPC and RCI suffer significant loss in sum rate when channel estimates are outdated due, for example, to mobile communication nodes or scatterers [4]. This work also demonstrated a heuristic beamformer that uses channel distribution information (CDI), in the form of spatial correlation matrices, to provide stable performance in the MIMO broadcast channel. Unlike DPC or RCI, this algorithm is able to use either CSI or CDI equally depending on the available information at the transmitter.

This paper adapts the RCI algorithm from [2] for use with CDI in the MIMO broadcast channel so that it provides stable throughput in the time-varying broadcast channel. The difficulty with this new algorithm, however, is that it requires the transmitter to know the full spatial correlation matrix, which for N transmit and receive antennas has N^4 matrix elements (compared to N^2 elements for the CSI). Consequently, it is desirable to reduce the volume of required feedback data without sacrificing throughput performance. This is accomplished in this paper by using channel models, namely the Kronecker [5] and Weichselberger [6] models, which parameterize the correlation matrix in an efficient manner.

2. Average Sum Rate Maximization

The $N \times 1$ vector of received signals for the jth user in the K- user broadcast channel can be written as

$$\mathbf{y}_j = \mathbf{H}_j \mathbf{b}_j x_j + \sum_{i \neq j}^{K} \mathbf{H}_j \mathbf{b}_i x_i + \boldsymbol{\eta}_j \tag{1}$$

where $\mathbf{H}_j$ is the $N \times N$ channel transfer matrix, $\mathbf{b}_j$ is the $N \times 1$ transmit beamformer, x_j is a unit variance Gaussian random variable representing the input signal, and $\boldsymbol{\eta}_j$ is additive white Gaussian noise (AWGN). Power is constrained such that $\sum_j \mathbf{b}_j^H \mathbf{b}_i = P$ where $\{\cdot\}^H$ is the matrix conjugate transpose. If the receiver applies a unit-length beamforming vector $\mathbf{w}_j$, the sum rate of the broadcast channel with linear processing can be written as

$$C = \sum_{j=1}^{K} \log\left(1 + \rho_j\right) \tag{2}$$

where

$$\rho_j = \frac{|\mathbf{w}_j^H \mathbf{H}_j \mathbf{b}_j|^2}{1 + \sum_{i \neq j} |\mathbf{w}_j^H \mathbf{H}_j \mathbf{b}_i|^2} \tag{3}$$

assuming unit variance noise and phase synchronization on the output signal.

The sum rate in Eq. (2) is maximized using beamforming weights chosen according to the RCI algorithm [2, 3] with perfect CSIT and CSIR, although the performance degrades rapidly as the CSI becomes outdated due to node motion [4]. Alternately, we can seek more stable performance by finding the beamformers that maximize a bound on the *average* sum capacity [4] or

$$\bar{C}_0 = \max_{\mathbf{w}_j, \mathbf{b}_j} \bar{C} = \max_{\mathbf{w}_j, \mathbf{b}_j} \sum_{j=1}^{K} \log\left(1 + \frac{\bar{n}_j}{\bar{d}_j}\right) \tag{4}$$

with power constraints on the transmit beamforming vectors, where $\bar{n}_j = \mathrm{E}[\mathrm{num}(\rho_j)]$, $\bar{d}_j = \mathrm{E}[\mathrm{den}(\rho_j)]$, and $\mathrm{num}(\cdot)$ and $\mathrm{den}(\cdot)$ return the numerator and denominator of the argument, respectively. Initiating this maximization by taking partial derivatives of $\bar{C}$ with respect to the elements of the transmit beamforming vectors produces weights $\mathbf{b}_j$ of the form

$$\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_K] = \left[\frac{\mathrm{Tr}(\mathbf{D})}{P}\mathbf{I} + \sum_{i=1}^{K} D_{ii} \overline{\mathbf{H}}_i\right]^{-1} \mathbf{\Lambda} \tag{5}$$

where $\mathrm{diag}(\cdot)$ returns a diagonal matrix of the inputs, $\mathrm{Tr}(\cdot)$ is the trace,

$$\begin{aligned}
\mathbf{D} &= \mathrm{diag}\left(\frac{\bar{n}_1}{\bar{d}_1(\bar{d}_1 + \bar{n}_1)}, \ldots, \frac{\bar{n}_K}{\bar{d}_K(\bar{d}_K + \bar{n}_K)}\right) \\
\overline{\mathbf{H}}_j &= \mathrm{mat}\left(\mathbf{S}_{\mathrm{t},j} \mathrm{vec}(\mathbf{w}_j \mathbf{w}_j^H)\right) \\
\mathbf{\Lambda} &= \left[\frac{(\overline{\mathbf{H}}_1 \mathbf{B})_1}{\bar{d}_1}, \ldots, \frac{(\overline{\mathbf{H}}_K \mathbf{B})_K}{\bar{d}_K}\right],
\end{aligned} \tag{6}$$

$\mathbf{S}_{\mathrm{t},j} = \mathrm{E}[\mathbf{H}_j^T \otimes \mathbf{H}_j^H]$, $\mathbf{S}_{\mathrm{r},j} = \mathrm{E}[\mathbf{H}_j^* \otimes \mathbf{H}_j]$, $\mathrm{vec}(\cdot)$ is the column stacking operator with inverse function $\mathrm{mat}(\cdot)$, $\{\cdot\}^T$ is the matrix transpose, and $\{\cdot\}^*$ is the conjugate. Using this notation, $\bar{n}_j = \mathbf{b}_j^H \overline{\mathbf{H}}_j \mathbf{b}_j$ and $\bar{d}_j = 1 + \sum_{i \neq j} \mathbf{b}_i^H \overline{\mathbf{H}}_j \mathbf{b}_i$.

The solution to this maximization is carried out by following the iterations suggested in [2, 3]. Following initialization of $\mathbf{D}$, $\mathbf{\Lambda}$, and $\mathbf{w}_j$, the algorithm computes $\overline{\mathbf{H}}_j$ and $\mathbf{B}$, allowing computation of new versions of $\mathbf{D}$ and $\mathbf{\Lambda}$. The minimum mean squared error (MMSE) criterion is then used to update $\mathbf{w}_j$ using $\mathbf{S}_{\mathrm{r},j}$ [4]. This process is repeated until convergence. This regularized channel distribution inversion (RCDI) beamformer reduces exactly to the RCI beamformer when the expectation operator is removed (i.e. $\mathbf{S}_{\mathrm{t},j} = \mathbf{H}_j^T \otimes \mathbf{H}_j^H$) and for the same initial conditions.

An important dependence of the RCI and RCDI beamformers are the initial conditions $\mathbf{D}$ and $\mathbf{\Lambda}$. Due to the nonconvex nature of the beamforming capacity expression, the algorithms only guarantee convergence to a local maximum and may not produce the true sum capacity of the broadcast channel with linear precoding [2], though a good starting point for RCI is the regularized pseudo-inverse of the channel. Because an analogous initial condition for the RCDI algorithm has not yet been derived, several different starting points are used and the selected solution is the one which produces the highest lower bound on average sum rate.

3. Channel Distribution Parameterization

The RCDI algorithm requires input parameters $\mathbf{S}_{\mathrm{t}}$ and $\mathbf{S}_{\mathrm{r}}$ that are nonlinear permutations on the large spatial correlation matrix $\mathbf{R} = \mathrm{E}[\mathrm{vec}(\mathbf{H})\mathrm{vec}(\mathbf{H})^H]$, where the user index is dropped for convenience. This section details how to use various channel models to parameterize $\mathbf{S}_{\mathrm{t}}$ and $\mathbf{S}_{\mathrm{r}}$ and limit the amount of required feedback. The following development uses $\mathbf{I}_{\mathrm{t}} = \mathrm{E}[\mathbf{H}_{\mathrm{w}}^T \otimes \mathbf{H}_{\mathrm{w}}^H]$ and $\mathbf{I}_{\mathrm{r}} = \mathrm{E}[\mathbf{H}_{\mathrm{w}}^* \otimes \mathbf{H}_{\mathrm{w}}]$, where $\mathbf{H}_{\mathrm{w}}$ is an $N \times N$ matrix with zero-mean, unit variance, i.i.d. complex Gaussian entries. Additionally, two properties of the Kronecker product are used: $\mathbf{AB} \otimes \mathbf{CD} = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$ and $(\mathbf{A} \odot \mathbf{B}) \otimes (\mathbf{C} \odot \mathbf{D}) = (\mathbf{A} \otimes \mathbf{C}) \odot (\mathbf{C} \otimes \mathbf{D})$ where $\odot$ is the matrix element-by-element product operator.

The **correlation model** uses the full spatial correlation matrix to model behavior in the channel. Under this model, channel realizations are generated using

$$\mathbf{H}_{\text{Corr}} = \text{mat}\left\{\sqrt{\mathbf{R}}\text{vec}(\mathbf{H}_{\text{w}})\right\}.$$

Given this model, the input parameters for the RCDI algorithm can be written as

$$\begin{aligned}\mathbf{S}_{\text{t}}^{\text{Corr}} &= \text{E}[\mathbf{H}_{\text{Corr}}^{T} \otimes \mathbf{H}_{\text{Corr}}^{H}] \\ \mathbf{S}_{\text{r}}^{\text{Corr}} &= \text{E}[\mathbf{H}_{\text{Corr}}^{*} \otimes \mathbf{H}_{\text{Corr}}]\end{aligned} \tag{7}$$

which are only functions of $\mathbf{R}$. For the correlation model the number of complex valued numbers needed for feedback is N^4.

The **Kronecker model** [5] assumes separability between transmit and receive correlation matrices. This decoupling of transmit and receive antennas allows channel realizations to be generated by

$$\mathbf{H}_{\text{Kron}} = \sqrt{\mathbf{R}_{\text{r}}}\mathbf{H}_{\text{w}}\sqrt{\mathbf{R}_{\text{t}}}$$

where the one-sided correlation matrices are calculated from $\mathbf{R}_{\text{r}} = \text{E}[\mathbf{H}\mathbf{H}^H]$ and $\mathbf{R}_{\text{t}} = \text{E}[\mathbf{H}^H\mathbf{H}]$. The input parameters to the RCDI algorithm assuming the Kronecker model reduce to

$$\begin{aligned}\mathbf{S}_{\text{t}}^{\text{Kron}} &= \left(\sqrt{\mathbf{R}_{\text{t}}}^{T} \otimes \sqrt{\mathbf{R}_{\text{t}}}^{H}\right)\mathbf{I}_{\text{t}}\left(\sqrt{\mathbf{R}_{\text{r}}}^{T} \otimes \sqrt{\mathbf{R}_{\text{r}}}^{H}\right) \\ \mathbf{S}_{\text{r}}^{\text{Kron}} &= \left(\sqrt{\mathbf{R}_{\text{r}}}^{*} \otimes \sqrt{\mathbf{R}_{\text{r}}}\right)\mathbf{I}_{\text{r}}\left(\sqrt{\mathbf{R}_{\text{t}}}^{*} \otimes \sqrt{\mathbf{R}_{\text{t}}}\right).\end{aligned} \tag{8}$$

For this work, we first decompose the full correlation matrix using a rank-1 approximation to the Kronecker product [7] which determines the one-sided correlations that minimize $||\mathbf{R} - \widehat{\mathbf{R}}_{\text{r}} \otimes \widehat{\mathbf{R}}_{\text{t}}||^2$. The estimates $\widehat{\mathbf{R}}_{\text{r}}$, $\widehat{\mathbf{R}}_{\text{t}}$ are then fed back to the transmitter and used to construct the necessary matrix estimates in Eq. (8). The complexity in the feedback channel is $2N^2$ complex numbers.

The **Weichselberger model** [6] uses a coupling matrix to model the "cross"-correlations between transmit and receive antennas. For this model, channel realizations are generated with

$$\mathbf{H}_{\text{Weichs}} = \mathbf{U}_{\text{r}}(\widetilde{\mathbf{\Omega}} \odot \mathbf{H}_{\text{w}})\mathbf{U}_{\text{t}}^{T}$$

where $\widetilde{\mathbf{A}}$ is the element-wise square root on the matrix $\mathbf{A}$ and the matrices $\mathbf{U}_{\text{r}}$ and $\mathbf{U}_{\text{t}}$ contain the eigenvectors of $\mathbf{R}_{\text{r}}$ and $\mathbf{R}_{\text{t}}$ from the Kronecker model, respectively. Using properties of the Kronecker product

$$\begin{aligned}\mathbf{S}_{\text{t}}^{\text{Weichs}} &= (\mathbf{U}_{\text{t}} \otimes \mathbf{U}_{\text{t}}^{*})\left\{(\widetilde{\mathbf{\Omega}}^{T} \otimes \widetilde{\mathbf{\Omega}}) \odot \mathbf{I}_{\text{t}}\right\}\left(\mathbf{U}_{\text{r}}^{T} \otimes \mathbf{U}_{\text{r}}^{H}\right) \\ \mathbf{S}_{\text{r}}^{\text{Weichs}} &= (\mathbf{U}_{\text{r}}^{*} \otimes \mathbf{U}_{\text{r}})\left\{(\widetilde{\mathbf{\Omega}}^{*} \otimes \widetilde{\mathbf{\Omega}}) \odot \mathbf{I}_{\text{r}}\right\}\left(\mathbf{U}_{\text{t}}^{H} \otimes \mathbf{U}_{\text{t}}^{T}\right).\end{aligned} \tag{9}$$

The Weichselberger model requires a complexity of $3N^2$ in the feedback channel.

4. Results

To show the performance of the suggested beamformer we first use the full correlation matrix model to produce estimates of $\mathbf{S}_{\text{t}}$ and $\mathbf{S}_{\text{r}}$. Channel realizations are obtained from an experimental sounder developed at Brigham Young University [8]. The dataset used in these results was taken at a center frequency of 2.55 GHz in a typical outdoor campus environment consisting of buildings, pedestrians, and light foliage.

Figure 1 shows complimentary cumulative distribution functions (CCDF) for the sum rate of the RCI (CSIT) and RCDI (CDIT) algorithms. The simulations use the measured data with $N = K = 4$ and a power of $P = 10$. A virtual delay (measured in wavelengths) is created in the feedback channel to the transmitter while the receiver has delay-free CSI. The RCDI beamformer is resilient to delays in the feedback channel while RCI degrades rapidly.

The effects of channel distribution parameterization are examined in Fig. 2. For this plot, the power was held constant at $P = 10$ while the number of antennas and users was swept. The results show that parameterizing the channel distribution does not significantly affect the average sum rate despite the significant savings in feedback complexity. For $N = 6$ the total number of complex numbers fed back per user for the correlation model is 1296 compared to only 108 and 72 for the Weichselberger and Kronecker models, respectively - an order of magnitude

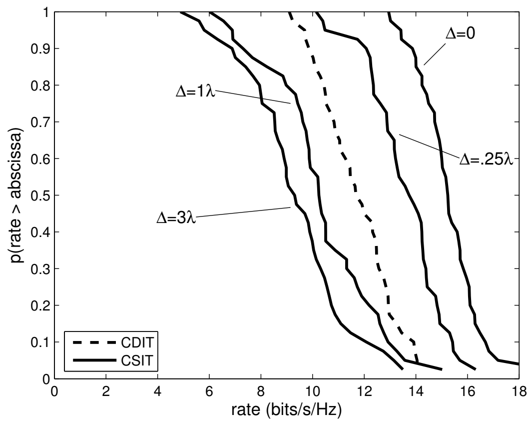

Fig. 1. Rate CCDF for the measured broadcast channel with $N = K = 4$ and $P = 10$ for RCDI using either CSI or CDI. The transmitter uses CSI outdated by Δ (wavelengths) from the current position.

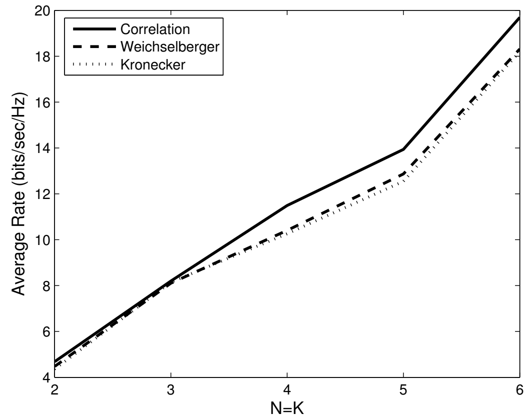

Fig. 2. Average rate versus system size for $\mathbf{S}_{\mathrm{t}}$ and $\mathbf{S}_{\mathrm{r}}$ generated by the correlation, Kronecker, and Weichselberger models.

difference. It should be reiterated that the models are only used to parameterize the experimentally-obtained channel for feedback and computation of the weights and not to generate the channels used in the simulations.

5. Conclusion

The sum rate maximizing beamformer has been adapted to use CDI as well as CSI depending on available channel information. This beamforming algorithm is robust to temporal variations and delay in the feedback channel. Furthermore, simple parameterization of the channel correlation matrix using well known channel models enables a significant reduction in the amount of information fed back to the transmitter while only resulting in a small loss in throughput performance.

Acknowledgements: This work was supported by the U. S. Army Research Office under the Multi-University Research Initiative (MURI) Grants # W911NF-04-1-0224 and # W911NF-07-1-0318.

6. References

[1] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Selected Areas Commun.*, vol. 21, pp. 684–702, June 2003.

[2] M. Stojnic, H. Vikalo, and B. Hassibi, "Rate maximization in multi-antenna broadcast channels with linear preprocessing," *IEEE Trans. Commun.*, vol. 5, pp. 2338–2342, Sept. 2006.

[3] Q. H. Spencer, J. W. Wallace, C. B. Peel, T. Svantesson, A. L. Swindlehurst, and A. Gummalla, "Performance of multi-user spatial multiplexing with measured channel data," in *MIMO System Technology and Wireless Communications*. CRC, 2006.

[4] A. Anderson, J. R. Zeidler, and M. A. Jensen, "Stable transmission in the time-varying MIMO broadcast channel," to appear in *EURASIP Journal on Advances in Sig. Proc.*

[5] S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Khan, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

[6] W. Weichselberger, M. Herdin, H. Özcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, no. 1, pp. 90–99, 2006.

[7] C. F. V. Loan and N. Pitsianis, "Approximations with Kronecker products," in *Linear Algebra for Large Scale and Real Time Applications*, M. S. Moonen and G. H. Golub, Eds. Kluwer Academic Publishers, 1993.

[8] J. Wallace and M. Jensen, "Time varying MIMO channels: Measurement, analysis, and modeling," *IEEE Trans. Antennas Propag.*, vol. 54, pp. 3265–3273, Nov. 2006.