

NON-ROBUSTNESS OF STATISTICS-BASED BEAMFORMER DESIGN IN CORRELATED MIMO CHANNELS

Vasanthan Raghavan, Ada S. Y. Poon, and Venugopal V. Veeravalli

ABSTRACT

Recent attention on correlated multi-input multi-output systems has centered around the case of imperfect channel or statistical information at the transmitter. The focus of this work is on correlated channels with arbitrary antenna array geometry and spacings, a coherent receiver, and imperfect statistical knowledge at the transmitter. Leveraging a recently proposed channel modeling paradigm that exploits processing in the *angular* domain, we first elucidate the structure of the optimal schemes when ‘genie-aided’ perfect statistical information is available at the transmitter. In the low-SNR case, we then show that the beamforming scheme that is optimal when perfect statistical information is available does not *degrade smoothly* with imperfections in the statistical information. We then go on to show that there exists a certain low-complexity beamforming approach which, while being sub-optimal in the genie-aided case, is *robust* to statistical feedback as well as the dynamics of its evolution.

Index Terms— Arbitrary antenna array architectures, beamforming, correlation, fading channels, imperfect channel information, MIMO systems, nonuniform sampling, robustness.

1. INTRODUCTION

Research over the last decade has firmly established the utility of multiple antennas at the transmitter and the receiver in achieving significant spectral efficiency gains. The realizability of such gains in the presence of practical impairments like spatial correlation, imperfect channel/statistical information at the transmitter etc. has been the subject of much recent attention. In this direction, many statistical channel models (applicable under different restrictions on the scattering environments) have been proposed in the literature and performance metrics such as the information theoretic capacity, error probability etc. have been studied.

Recently in [1], we proposed a channel model that is applicable for antenna arrays with arbitrary geometry and antenna spacings and is a generalization of the virtual representation [2] for uniform linear arrays (ULAs). Furthermore, we characterized the capacity of such channels with a coherent receiver and ‘genie-aided’ perfect statistical information at the transmitter; see Lemma 1 below. We also showed that beamforming along a fixed (but appropriately transformed) direction is optimal in the low-SNR extreme, while uniform

The authors are with the Coordinated Science Laboratory and the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA. Email: {vasanth, poon, vvv}@uiuc.edu.

power signaling over a fixed (but appropriately transformed) eigen-space is optimal in the high-SNR extreme (Lemma 2). The focus of this work is on the robustness of the low-SNR optimal scheme to imperfections in statistical information. Assuming that statistical information is fed back to the transmitter via a B -bit error-free quantization scheme, we show that for any fixed choice of B , there exist scattering environments that show a significant degradation in performance with statistical uncertainty. On the other hand, we show that a low-complexity scheme that beamforms along the dominant eigen-mode of the transmit covariance matrix (which is sub-optimal in the genie-aided case) is more robust to uncertainties in statistics. The loss in performance with this scheme depends only on how close the transmit spatial array matrix is to being unitary and is (more or less) independent of the value of B .

2. SYSTEM MODEL

Our focus is on a multi-antenna communication system with N_t transmitters and N_r receivers and the system equation

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

where \mathbf{y} is the $N_r \times 1$ received vector, \mathbf{x} is the $N_t \times 1$ transmitted vector, and \mathbf{n} is the $N_r \times 1$ additive white Gaussian noise vector. The channel matrix \mathbf{H} can be written as

$$\mathbf{H} = \int \int \int h(t, k_r, k_t) e^{-j2\pi tW} \mathbf{a}_r(k_r) \mathbf{a}_t(k_t)^H dt dk_r dk_t$$

where $h(t, k_r, k_t)$ corresponds to the zero-mean complex gain between transmit direction¹ k_t and receive direction k_r at time t due to an impulse applied at time 0, W is the signaling bandwidth, and $\mathbf{a}_t(k_t)$ and $\mathbf{a}_r(k_r)$ denote the $N_t \times 1$ transmit and $N_r \times 1$ receive array steering vectors, respectively.

While the above equation corresponds to a continuous channel representation [3], the system equation can be discretized by exploiting the following fact [1]: For a given antenna array geometry and spacings, there exist matrices \mathbf{A}_t and \mathbf{A}_r such that the channel \mathbf{H} can be efficiently approximated as

$$\mathbf{H} \approx \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} \mathbf{H}_{\text{ang}}[i, j] \mathbf{a}_r(i) \mathbf{a}_t^H(j) = \mathbf{A}_r \mathbf{H}_{\text{ang}} \mathbf{A}_t^H \quad (2)$$

where $\mathbf{H}_{\text{ang}} = \{\mathbf{H}_{\text{ang}}[i, j]\}$ denotes the channel in the *angular* domain with $\mathbf{H}_{\text{ang}}[i, j]$ approximately independent with

¹Both directions measured with respect to the horizontal axes, without any loss in generality.

zero mean, and $\mathbf{a}_t(i)$ and $\mathbf{a}_r(j)$ denote the i -th and j -th column vectors of \mathbf{A}_t and \mathbf{A}_r , respectively. Assuming sufficiently rich scattering, the channel coefficients can be modeled as Gaussian and hence, the statistics of \mathbf{H}_{ang} are determined by the second moments of its entries.

Critical to (2) is the existence of an appropriate, but fixed non-uniform sampling in the physical domain so as to create a uniform, nonoverlapping partition in the *angular* domain (this is analogous to the non-uniform sampling in time considered in [4]). The fixed linear transformation matrices that achieve this sampling in the transmit and the receive domain are denoted by \mathbf{A}_t and \mathbf{A}_r , respectively. By definition, these matrices are non-singular; however, they may not be unitary or even have unit-normed row or column vectors. In the special case of uniform linear arrays (ULAs) at both the ends, \mathbf{A}_t and \mathbf{A}_r reduce to unitary discrete Fourier transform matrices [2]. Note that while it may be possible to perform a Karhunen-Loeve type decomposition of the channel \mathbf{H} as in (2) with a scattering environment-dependent choice of \mathbf{A}_t and \mathbf{A}_r as in [5], it is crucial to note that the choices of \mathbf{A}_t and \mathbf{A}_r that we propose in (2) are in fact independent of the scattering environment and are dependent only on the antenna array geometry and spacing. In this aspect, our work follows the philosophy prescribed in [2, 6]. The non-unitarity of \mathbf{A}_t and \mathbf{A}_r lead to fundamentally new issues that will be explored in this paper.

3. PERFORMANCE OF THE OPTIMAL GENIE-AIDED SCHEME

In this section, we assume that perfect channel information is available at the receiver while perfect (array and statistical) information, that is, \mathbf{A}_t , \mathbf{A}_r and $\sigma_{ij}^2 = \mathbf{E}[|\mathbf{H}_{\text{ang}}[i, j]|^2]$ are available at the transmitter. We call this scheme ‘genie-aided’ and distinguish it from a practical (low-complexity) scheme (with imperfect channel information) which is the focus of the subsequent sections. The performance of the genie-aided scheme serves as a theoretical benchmark to compare the performance of the practical scheme.

In this setting, the capacity C of the MIMO system with spatial array matrices $\{\mathbf{A}_t, \mathbf{A}_r\}$ is

$$\begin{aligned} C &= \sup_{\mathbf{Q} \in \mathcal{Q}} \mathbf{E} [\log_2 \det (\mathbf{I}_{N_r} + \mathbf{A}_r \mathbf{H}_{\text{ang}} \mathbf{A}_t^H \mathbf{Q} \mathbf{A}_t \mathbf{H}_{\text{ang}}^H \mathbf{A}_r^H)] \\ &= \sup_{\mathbf{Q} \in \mathcal{Q}} \mathbf{E} [\log_2 \det ((\mathbf{A}_r^H \mathbf{A}_r)^{-1} + \mathbf{H}_{\text{ang}} \mathbf{A}_t^H \mathbf{Q} \mathbf{A}_t \mathbf{H}_{\text{ang}}^H)] \\ &\quad + \log_2 \det (\mathbf{A}_r^H \mathbf{A}_r) \end{aligned}$$

where the optimization is over the set \mathcal{Q} defined as $\mathcal{Q} = \{\mathbf{Q} : \mathbf{Q} \succ \mathbf{0}, \text{Tr}(\mathbf{Q}) \leq \rho\}$. In prior work [1], we addressed the issue of optimal signaling, summarized² below.

Lemma 1 *Given that the transmitter knows \mathbf{A}_t , \mathbf{A}_r and the statistics of \mathbf{H}_{ang} , the optimal input is zero mean Gaussian with input covariance matrix given by $\mathbf{Q}_{\text{opt}} = \mathbf{A}_t^{-H} \mathbf{D}_{\text{opt}} \mathbf{A}_t^{-1}$ where \mathbf{D}_{opt} is a positive semi-definite, diagonal matrix that solves the following convex optimization problem:*

$$\mathbf{D}_{\text{opt}} = \arg \max_{\mathbf{D}} \mathbf{E} [\log_2 \det ((\mathbf{A}_r^H \mathbf{A}_r)^{-1} + \mathbf{H}_{\text{ang}} \mathbf{D} \mathbf{H}_{\text{ang}}^H)]$$

such that $\text{Tr}(\mathbf{D}(\mathbf{A}_t^H \mathbf{A}_t)^{-1}) \leq \rho$.

²See [1] for proofs of all statements.

The above optimization problem is a very standard problem: maximizing a concave function over a convex set and hence many of the standard convex optimization procedures are applicable. In the special case of ULAs [6], it is known that in the low-SNR extreme, beamforming to the statistically dominant transmit eigen-direction is capacity optimal while in the high-SNR extreme, uniform power signaling across the rank of \mathbf{H}_{ang} is capacity optimal. More generally, in the intermediate-SNR regime, the rank of \mathbf{Q}_{opt} is non-decreasing as ρ increases. With the more general channel model, these results extend as follows.

Lemma 2 *Define the direction j^* as*

$$j^* = \arg \max_j \frac{\sum_i \sigma_{ij}^2 (\sum_k |\mathbf{A}_r[k, i]|^2)}{(\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j]}. \quad (3)$$

Beamforming along j^ (with $\mathbf{Q} = \mathbf{A}_t^{-H} \mathbf{D} \mathbf{A}_t^{-1}$, $\text{Tr}(\mathbf{Q}) = \rho$, and \mathbf{D} having only the diagonal entry in the j^* -th position non-zero) is optimal in the low-SNR extreme. Let \mathcal{R} denote the set of indices $\{j : \sum_i \sigma_{ij}^2 (\sum_k |\mathbf{A}_r[k, i]|^2) > 0\}$. Uniform power signaling scheme across \mathcal{R} is optimal in the high-SNR extreme.*

4. PERFORMANCE OF A LOW-COMPLEXITY SCHEME

We now focus on the low-SNR extreme and establish connections between beamforming in the ULAs case and the non-ULAs case. Consider the case where \mathbf{A}_t and \mathbf{A}_r are unitary. In this setting, \mathbf{Q}_{opt} reduces to $\mathbf{A}_t \mathbf{D}_{\text{opt}} \mathbf{A}_t^H$ (with the beamforming direction $j^* = \arg \max_j \sum_i \sigma_{ij}^2$). That is, the optimal scheme is equivalent to beamforming along the dominant eigen-direction of the transmit covariance matrix (\mathbf{R}_t) defined as $\mathbf{R}_t \triangleq \mathbf{E} [\mathbf{H}^H \mathbf{H}]$. Note that \mathbf{R}_t is equal to $\mathbf{A}_t \mathbf{E} [\mathbf{H}_{\text{ang}}^H \mathbf{H}_{\text{ang}}] \mathbf{A}_t^H = \mathbf{A}_t \mathbf{D}_t \mathbf{A}_t^H$ with $\mathbf{D}_t[j] = \sum_i \sigma_{ij}^2$. This result is well-known, see e.g., [6].

It is natural to pose the question: How does this scheme perform in the non-ULA case? Towards answering this question, note the following relationship in the general case:

$$\begin{aligned} \mathbf{R}_t &= \mathbf{A}_t \mathbf{D}_t \mathbf{A}_t^H = \sum_{j=1}^{N_t} \mathbf{D}_t[j] \mathbf{a}_t(j) \mathbf{a}_t^H(j) \quad (4) \\ \mathbf{D}_t[j] &= \sum_i \sigma_{ij}^2 \sum_k |\mathbf{A}_r[k, i]|^2. \end{aligned}$$

Using the SVD of $\mathbf{A}_t (= \mathbf{U}_t \mathbf{\Lambda}_t \mathbf{V}_t^H)$, we have

$$\begin{aligned} \mathbf{R}_t &= \mathbf{U}_t (\mathbf{\Lambda}_t \mathbf{V}_t^H \mathbf{D}_t \mathbf{V}_t \mathbf{\Lambda}_t) \mathbf{U}_t^H \\ &= \mathbf{U}_t \widehat{\mathbf{U}} \widehat{\mathbf{\Lambda}} \widehat{\mathbf{U}}^H \mathbf{U}_t^H \quad (5) \end{aligned}$$

where we assume an SVD of the form $\widehat{\mathbf{U}} \widehat{\mathbf{\Lambda}} \widehat{\mathbf{U}}^H$ for the matrix in the parenthesis. We can also see that

$$\begin{aligned} \mathbf{Q}_{\text{opt}} &= \mathbf{A}_t^{-H} \mathbf{D}_{\text{opt}} \mathbf{A}_t^{-1} \\ &= \mathbf{U}_t (\mathbf{\Lambda}_t^{-1} \mathbf{V}_t^H \mathbf{D}_{\text{opt}} \mathbf{V}_t \mathbf{\Lambda}_t^{-1}) \mathbf{U}_t^H \\ &= \mathbf{U}_t \widetilde{\mathbf{U}} \widetilde{\mathbf{\Lambda}} \widetilde{\mathbf{U}}^H \mathbf{U}_t^H \quad (6) \end{aligned}$$

where an SVD of the form $\widetilde{\mathbf{U}} \widetilde{\mathbf{\Lambda}} \widetilde{\mathbf{U}}^H$ is assumed for the matrix in the parenthesis. From (5) and (6), we thus see that, in general the eigenvectors of \mathbf{Q}_{opt} and \mathbf{R}_t are not the same. In other

words, beamforming along the eigenvectors of \mathbf{R}_t is strictly sub-optimal from a capacity viewpoint in the non-ULA case. The following proposition captures the loss in performance that results through the use of this sub-optimal scheme. For this, we need the following definition.

Definition 1 Let \mathbf{B} be an $n \times n$ Hermitian matrix with

$$\delta_{\mathbf{B},j} = \frac{\left(\sum_{k \neq j} |\mathbf{B}[j, k]|\right)^{1-\alpha} \left(\sum_{k \neq j} |\mathbf{B}[k, j]|\right)^\alpha}{|\mathbf{B}[j, j]|}.$$

Then, \mathbf{B} is said to satisfy the Ostrowskii condition for some $\alpha \in (0, 1)$ if

$$\Delta_{\mathbf{B}} \triangleq \max_{j=1, \dots, n} \delta_{\mathbf{B},j} < 1. \quad (7)$$

The Ostrowskii condition is a measure of the diagonal dominance of \mathbf{B} : the more diagonally dominant \mathbf{B} is, the smaller $\Delta_{\mathbf{B}}$ is and *vice versa*. Note that the extreme cases of $\alpha = 0$ and $\alpha = 1$ correspond to whether the diagonal entries dominate the off-diagonal entries in the row or the column corresponding to that diagonal entry. The general case of $\alpha \in (0, 1)$ captures diagonal dominance over the weighted off-diagonal entries. A matrix satisfying the Ostrowskii condition is invertible.

Proposition 1 Let the average mutual information of the above sub-optimal scheme be $I_{\mathbf{R}_t, \text{bf}}(\rho)$ and consider the low-SNR regime, $\rho < \rho_{\text{low}}$ where ρ_{low} is defined as

$$\rho_{\text{low}} \triangleq \frac{(\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j^*]}{\sum_i \sigma_{ij}^2 (\sum_k |\mathbf{A}_r[k, i]|^2)}. \quad (8)$$

If $\mathbf{A}_t^H \mathbf{A}_t$ satisfies the Ostrowskii condition for some $\alpha \in (0, 1)$, we have (up to the dominant first-order term)

$$\frac{C(\rho) - I_{\mathbf{R}_t, \text{bf}}(\rho)}{C(\rho)} \leq \frac{(\Delta_{\mathbf{A}_t^H \mathbf{A}_t})^2}{1 - (\Delta_{\mathbf{A}_t^H \mathbf{A}_t})^2}. \quad (9)$$

The Ostrowskii condition on $\mathbf{A}_t^H \mathbf{A}_t$ implies that

$$\begin{aligned} \lambda_{\max}(\mathbf{A}_t^H \mathbf{A}_t) &\leq \max_j \left(\mathbf{A}_t^H \mathbf{A}_t[j] + \delta_{\mathbf{A}_t^H \mathbf{A}_t, j} \right) \\ &\leq 2 \max_j \mathbf{A}_t^H \mathbf{A}_t[j], \\ \lambda_{\min}(\mathbf{A}_t^H \mathbf{A}_t) &\geq \min_j \left(\mathbf{A}_t^H \mathbf{A}_t[j] - \delta_{\mathbf{A}_t^H \mathbf{A}_t, j} \right) > 0 \end{aligned}$$

and hence, \mathbf{A}_t is well-conditioned. In other words, the following general principle holds: As \mathbf{A}_t becomes more close to being unitary, $\mathbf{A}_t^H \mathbf{A}_t$ becomes more diagonally dominant, and hence, the value of $\Delta_{\mathbf{A}_t^H \mathbf{A}_t}$ reduces and the performance of the sub-optimal scheme approaches that of the optimal scheme. For example, (as noted earlier) in the special case where \mathbf{A}_t is unitary, both the schemes coincide and this conclusion is concurred by the first-order term in (9) reducing to zero.

5. ROBUSTNESS OF THE TWO SCHEMES

Most works in the literature assume that the channel statistics change at a sufficiently slow rate so that they can be acquired

at the transmitter with negligible cost. While this may be reasonable in certain low mobility situations, in general, the receiver has to estimate the statistical information and feed it back to the transmitter. We now account for the feedback cost and study the robustness of the two schemes to statistical feedback.

Since \mathbf{A}_t and \mathbf{A}_r are *independent* of the scattering environment, it is reasonable to assume that they can be acquired perfectly via either deterministic techniques upon knowledge of antenna geometry and spacings [4] or long-term averaging techniques that extract the array information by averaging out the imprint of the scattering environment in the short-term channel statistics. This knowledge implies that determining the optimal beamforming direction requires knowledge of only $\{\mathbf{D}_t[j] = \sum_i \sigma_{ij}^2 \sum_k |\mathbf{A}_r[k, i]|^2\}$ at the transmitter. Using its knowledge of \mathbf{A}_t , the transmitter then computes j^* as in (3). Thus, optimal beamforming requires feedback of N_t positive numbers, $\{\mathbf{D}_t[j]\}$, that are a function of the short-term channel statistics.

We assume the following B -bit quantization scheme for $\mathbf{D}_t[j]$ and error-free estimates $\widehat{\mathbf{D}}_t[j]$ at the transmitter. Let the transmitter and the receiver have an *a priori* knowledge of a quantization range $[0, \sigma_{\max}^2]$. For e.g., a suitable choice of σ_{\max}^2 is $\sigma_{\max}^2 = c \cdot \max_{\mathcal{E}} \max_j \sum_i \sigma_{ij}^2 \sum_k |\mathbf{A}_r[k, i]|^2$ for some suitable $c > 1$. The class \mathcal{E} corresponds to a representative family of scattering environments that the transmitter and the receiver encounter in a given communication scenario. The choice of c is such that for any scattering environment in \mathcal{E} and the associated statistics $\{\sigma_{ij}^2\}$, we have

$$\frac{\sigma_{\max}^2}{\sum_i \sigma_{ij}^2 \sum_k |\mathbf{A}_r[k, i]|^2} = \mathcal{O}(1). \quad (10)$$

The range $[0, \sigma_{\max}^2]$ is quantized with B -bits and hence the length of each quantization interval is $\frac{\sigma_{\max}^2}{2^B}$. By neglecting the error and delay associated with the feedback link, for any j , the error in estimating $\mathbf{D}_t[j]$ is at most $\delta_{\text{err}} = \frac{\sigma_{\max}^2}{2^{B+1}}$. The transmitter can use $\widehat{\mathbf{D}}_t[j]$ to compute the optimal beamforming direction akin to (3).

Alternatively, the transmitter could estimate \mathbf{R}_t as $\widehat{\mathbf{R}}_t = \sum_{j=1}^{N_t} \widehat{\mathbf{D}}_t[j] \mathbf{a}_t(j) \mathbf{a}_t^H(j)$. Then, the transmitter can signal sub-optimally along the dominant eigenvector of $\widehat{\mathbf{R}}_t$. The following results study the robustness of the two beamforming schemes to imperfections in statistical information. Given any fixed scattering environment, if B is large enough to quantize $\{\mathbf{D}_t[j]\}$ finely, the following proposition claims no loss due to statistical feedback.

Proposition 2 For any given scattering environment and the corresponding $\{\mathbf{D}_t[j]\}$, if B satisfies

$$\begin{aligned} B > \log_2 \left(\frac{1}{\min_{j \neq j^*} (\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j]} + \frac{1}{(\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j^*]} \right) \\ + \log_2 \left(\frac{\sigma_{\max}^2}{(\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j^*]} - \max_{j \neq j^*} \frac{\mathbf{D}_t[j]}{(\mathbf{A}_t^H \mathbf{A}_t)^{-1}[j]} \right) - 1 \triangleq B_{\min} \end{aligned}$$

then, the optimal scheme does not suffer any loss due to statistical feedback.

For a fixed environment, B_{\min} could be arbitrarily large. For example, consider \mathbf{A}_t unitary and $\mathbf{D}_t[j^*]$ arbitrarily close to $\max_{j \neq j^*} \mathbf{D}_t[j]$. That is, \mathbf{R}_t is such that its first two eigenvalues are close enough. Using the same logic as in Prop. 2, we have the following corollary.

Corollary 1 Let $I_{\text{bf},B}(\rho)$ denote the average mutual information of the optimal scheme with B -bit statistical feedback. There exist scattering environments and realizations $\{\sigma_{i,j}^2\}$ associated with them such that

$$\frac{C(\rho) - I_{\text{bf},B}(\rho)}{C(\rho)} \approx \mathcal{O}(1) \cdot \frac{1}{2^B}. \quad (11)$$

On the other hand, we have the following proposition.

Proposition 3 Let $I_{\mathbf{R}_t, \text{bf}, B}(\rho)$ denote the average mutual information of the sub-optimal scheme with B -bit statistical feedback. Let $\mathbf{U}_{\mathbf{R}_t}$ denote an eigenvector matrix of \mathbf{R}_t and denote the matrix $\mathbf{U}_{\mathbf{R}_t}^H \mathbf{A}_t$ by \mathbf{V} . Then, for any B , we have

$$\begin{aligned} & \frac{I_{\mathbf{R}_t, \text{bf}}(\rho) - I_{\mathbf{R}_t, \text{bf}, B}(\rho)}{C(\rho)} \\ & \leq \frac{\lambda_{\max}(\mathbf{A}_t^H \mathbf{A}_t)}{\lambda_{\min}(\mathbf{A}_t^H \mathbf{A}_t)} \cdot \sum_{i>1} \frac{|\sum_k \mathbf{V}[i, k](\widehat{\mathbf{D}}_t[k] - \mathbf{D}_t[k])\mathbf{V}^*[1, k]|^2}{(\lambda_{\max}(\mathbf{R}_t) - \lambda_i(\mathbf{R}_t))^2}. \end{aligned}$$

Note that \mathbf{V} is the projection of the eigenvectors of \mathbf{R}_t on to the transmit spatial array matrix. When \mathbf{A}_t is unitary, \mathbf{V} reduces to \mathbf{I} and the above upper bound reduces to zero since the summation is over the indices $i > 1$. Similarly, when \mathbf{A}_t is close to being unitary, \mathbf{V} is close to being \mathbf{I} and the upper bound is small even for $B = 1$. That is, the performance of the sub-optimal scheme is critically dependent on how close \mathbf{A}_t is to being unitary with very little dependence on the choice of B . Now, we can combine Props. 1- 3 and Cor. 1 to obtain the main statement of this paper.

Theorem 1 For any realization of the channel statistics, the relative loss in average mutual information of the sub-optimal scheme satisfies

$$\begin{aligned} & \frac{C(\rho) - I_{\mathbf{R}_t, \text{bf}, B}(\rho)}{C(\rho)} \leq \frac{(\Delta_{\mathbf{A}_t^H \mathbf{A}_t})^2}{1 - (\Delta_{\mathbf{A}_t^H \mathbf{A}_t})^2} \\ & + \frac{\lambda_{\max}(\mathbf{A}_t^H \mathbf{A}_t)}{\lambda_{\min}(\mathbf{A}_t^H \mathbf{A}_t)} \cdot \sum_{i>1} \frac{|\sum_k \mathbf{V}[i, k](\widehat{\mathbf{D}}_t[k] - \mathbf{D}_t[k])\mathbf{V}^*[1, k]|^2}{(\lambda_{\max}(\mathbf{R}_t) - \lambda_i(\mathbf{R}_t))^2}. \end{aligned}$$

If $B > B_{\min}$ (which is given in Prop. 2), then the optimal scheme incurs no loss due to statistical feedback. However, for any fixed B , there exists channel statistics such that the relative loss in average mutual information of the optimal scheme satisfies

$$\frac{C(\rho) - I_{\text{bf},B}(\rho)}{C(\rho)} \approx \frac{\mathcal{O}(1)}{2^B}.$$

6. DISCUSSION

We now focus attention on the practical case of small B (corresponding to a low-rate feedback scheme). Our main conclusion is: *Given that \mathbf{A}_t is known at the transmitter (and is*

unitary), statistical feedback has negligible impact on the sub-optimal scheme. On the other hand, for any given scattering environment, if $B < B_{\min}$ (corresponding to that environment), the sub-optimal scheme *can* lead to a higher throughput than the optimal³ scheme.

However, more important is the behavior of the optimal scheme in response to change in channel statistics (possibly due to mobility in realistic situations). Cor. 1 establishes that the low-SNR optimal beamforming scheme is very sensitive to feedback in this regime and suffers a worst-case relative distortion on the order of 2^{-B} . In the small- B regime, this relative loss could be significantly larger than that incurred by the sub-optimal scheme (with feedback). Thus, from a robustness perspective, the sub-optimal scheme can perform much better than the optimal scheme in practice. The advantage of the sub-optimal scheme in the case of statistical feedback is also intuitive upon closer introspection of the structure of the two signaling schemes. Note that the optimal signaling scheme performs a transmit array (transposition and) inversion which reduces to signaling along the transmit array manifold when \mathbf{A}_t is near-unitary. In this case, even with low levels of feedback (small- B), the eigenvectors of $\widehat{\mathbf{R}}_t$ are close to those of \mathbf{A}_t .

The above discussion suggests that there is a need to reinvestigate the structure and the robustness of optimal signaling schemes with practical impairments such as channel state feedback. In the final version of this paper, we will provide insights on the robustness of the performance of the optimal rank- M scheme ($M > 1$) with respect to the sub-optimal scheme that excites the M dominant eigenmodes of \mathbf{R}_t . We will also provide numerical results to illustrate our results.

7. REFERENCES

- [1] V. Raghavan, A. S. Y. Poon, and V. V. Veeravalli, "Angular Domain Processing for MIMO Systems: Capacity and Low-Complexity Schemes," *To appear, Proc. Asilomar Conf.*, 2007, Available: [Online] <http://www.ifp.uiuc.edu/~vasanth>.
- [2] A. M. Sayeed, "Deconstructing Multi-Antenna Fading Channels," *IEEE Trans. Sig. Proc.*, vol. 50, no. 10, pp. 2563–2579, Oct. 2002.
- [3] A. S. Y. Poon, R. W. Brodersen, and D. N. C. Tse, "Degrees of Freedom in Multiple-Antenna Channels: A Signal Space Approach," *IEEE Trans. Inform. Theory*, vol. 51, no. 2, pp. 523–536, Feb. 2005.
- [4] R. Venkataramani and Y. Bresler, "Optimal Sub-Nyquist Nonuniform Sampling and Reconstruction for Multiband Signals," *IEEE Trans. Sig. Proc.*, vol. 49, no. 10, pp. 2301–2313, Oct. 2001.
- [5] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Canonical Statistical Modeling and Capacity Analysis of Correlated MIMO Fading Channels," *Submitted to IEEE Trans. Inform. Theory*, Sept. 2007, Available: [Online] <http://www.ifp.uiuc.edu/~vasanth>.
- [6] V. V. Veeravalli, Y. Liang, and A. M. Sayeed, "Correlated MIMO Rayleigh Fading Channels: Capacity, Optimal Signaling and Asymptotics," *IEEE Trans. Inform. Theory*, vol. 51, no. 6, pp. 2058–2072, June 2005.

³Optimal and sub-optimal from a genie-aided statistics acquisition perspective.