Structured Statistical Precoding for Correlated MIMO Channels

Vasanthan Raghavan, Akbar M. Sayeed, Venugopal V. Veeravalli
vasanth@uiuc.edu, akbar@engr.wisc.edu, vvv@uiuc.edu

***Abstract*— The focus of this paper is on spatial precoding in correlated multi-antenna channels where the number of independent data-streams can be adapted to trade-off the data-rate with the transmitter complexity. A *structured precoding* scheme is proposed, where the precoder structure evolves fairly slowly at a rate comparable with the statistical evolution of the channel, and in addition, enjoys low-complexity. A particular case of the proposed scheme, semiunitary precoding, is shown to be near-optimal in *matched channels* where the dominant eigenvalues of the transmit covariance matrix are well-conditioned and their number equals the number of independent data-streams, and the receive covariance matrix is also well-conditioned. In *mismatched channels* where the above conditions do not hold, it is shown that the loss in performance with semiunitary precoding when compared with a perfect channel information benchmark is substantial. This loss can be mitigated via limited feedback techniques that provide partial channel information to the transmitter. We also develop matching metrics that capture the degree of matching of a channel to the precoder structure *continuously*, and allow ordering two matrix channels in terms of their mutual information or error probability performance.**

I. Introduction

Multiple antenna communication has received significant attention over the last decade as a mechanism to increase reliability, or data rates, or a combination of the two. The focus of this work is on point-to-point spatial precoding systems where the number of independent data-streams is constrained to be a subset, M, of the transmit dimension so as to minimize the complexity and the cost associated with transmission. Initial works study optimal signaling strategies when perfect channel state information (CSI) is available at the transmitter and the receiver. These studies show that a *channel diagonalizing* input that excites the dominant M-dimensional eigen-space of the channel, with a power allocation that is computed via waterfilling, is robust under different design metrics (see [1]–[4] and references therein).

In practical systems, where the channels are spatially correlated, it is reasonable to assume statistical knowledge at the transmitter. Optimal precoding with the statistical information alone has also been studied extensively (see [5]–[7] etc.). In this case, the eigen-directions of the optimal input covariance matrix correspond to a set of M-dominant eigenvectors of the transmit covariance matrix and are hence, easily adaptable to changes in statistics. However, computing the power allocation across the M modes requires Monte Carlo averaging or gradient descent approaches, the complexities of which are questionable under certain situations (e.g., at the mobile ends). More importantly, it is not clear as to how far the performance of the statistical scheme is with respect to the perfect CSI benchmark.

It should be noted that the above works study precoder design with an emphasis on obtaining information-theoretic limits on performance. In contrast, our goal here is on low-complexity schemes that can be easily implemented and adapted to changes in channel statistics. Towards this goal, we consider a narrowband setup where spatial correlation is modeled by a general channel decomposition, based on physical principles, that has been verified by many recent measurement campaigns, and includes as special cases many well-studied channel models.

We propose the notion of *structured precoding* where the power allocation across the M modes is either fixed and known at both the ends or computable with low-complexity, possibly as a function of the signal-to-noise ratio (SNR) and the statistics. In particular, we study the performance of a statistical semiunitary precoding scheme where the eigen-directions of the input correspond to the dominant M-dimensional eigen-space of the transmit covariance matrix and the power allocation across the modes is equal. Our focus is on two questions: 1) When is the semiunitary scheme near-optimal with respect to a perfect CSI benchmark?, and 2) If it is not near-optimal, what is the gap in performance that can possibly be bridged with a *limited feedback* scheme [8] that provides partial channel information to the transmitter?

The answers to the above questions lie in the notion of *matched* and *mismatched channels*, which are introduced in this work. A matched channel is one where the channel is effectively matched to the precoding scheme with the following two conditioning properties being true: 1) The M-dominant eigenvalues of the transmit covariance matrix are *well-conditioned*[1], whereas the remaining $(N_t - M)$ eigenvalues are *ill-conditioned* away from the dominant ones, and 2) The receive covariance matrix is also *well-conditioned*. A mismatched channel is one where both the transmit and the receive covariance matrices are ill-conditioned, with the additional condition that $\mathsf{rank}(\mathbf{H}) \geq M$.

We show that matched and mismatched channels correspond to the cases where the relative performance of the semiunitary precoder are closest and farthest to the perfect CSI precoder, respectively. The degree of channel-to-precoder scheme matching can be abstractly measured with *matching metrics*, that are also introduced in this work. As a by-product of our study, we also show that the statistical precoder is near-optimal in the *relative antenna asymptotic setting*[2] for any channel. This conclusion generalizes our previous work in the beamforming case ($M = 1$), where we studied the reliability performance of the statistical beamforming scheme [9]. Proofs of all statements in this paper can be found in [10].

II. System Setup

We consider a communication system with N_t transmit and N_r receive antennas where M ($1 \leq M \leq N_t$) independent data-streams are used in signaling. The baseband model is

$$\mathbf{y} = \sqrt{\frac{\rho}{M}}\,\mathbf{H}\,\mathbf{F}\,\mathbf{s} + \mathbf{n} \tag{1}$$

[1] If $\boldsymbol{\Lambda}_t(1) \geq \cdots \geq \boldsymbol{\Lambda}_t(M)$ denote the first M eigenvalues of the transmit covariance matrix and $\frac{\boldsymbol{\Lambda}_t(1)}{\boldsymbol{\Lambda}_t(M)}$ is (or is not) significantly larger than 1, we loosely say that these eigenvalues are ill-(or well-)conditioned.

[2] That is, when $\frac{M}{N_r} \rightarrow 0$ or ∞ as $\{M, N_t, N_r\} \rightarrow \infty$.

where ρ is the transmit power constraint, the $M \times 1$ input vector $\mathbf{s}$ has i.i.d. zero mean unit variance components, $\mathbf{F}$ is the $N_t \times M$ precoder with $\mathrm{Tr}(\mathbf{F}^H \mathbf{F}) \leq M$ and $\mathbf{n}$ is the N_r-dimensional standard additive white Gaussian noise. Without loss in generality, it can be assumed that $\mathbf{F} = \mathbf{V_F} \, \mathbf{\Lambda_F}^{1/2}$ where $\mathbf{V_F}$ is an $N_t \times M$ semiunitary[3] matrix and $\mathbf{\Lambda_F}$ is an $M \times M$ positive-definite power-shaping matrix with $\mathrm{Tr}(\mathbf{\Lambda_F}) \leq M$.

We assume a block fading, narrowband model for the correlation of the channel in time and frequency. The main emphasis in this work is on the impact of correlation in the spatial domain. We assume a Rayleigh fading model for the channel coefficients. The second order statistics are described via a general, mathematically tractable *canonical decomposition* of the channel along the transmit and the receive covariance bases [11], [12]. In this model, we assume that the auto- and the cross-covariance matrices of all the rows and columns of $\mathbf{H}$ share the same eigen-bases. Thus, $\mathbf{H}$ is decomposed as

$$\mathbf{H} = \mathbf{U}_r \, \mathbf{H}_{\mathsf{ind}} \, \mathbf{U}_t^H \tag{2}$$

where $\mathbf{H}_{\mathsf{ind}}$ has independent, but not necessarily identically distributed entries, and $\mathbf{U}_t$ and $\mathbf{U}_r$ are unitary matrices. The transmit and the receive covariance matrices are given by

$$\begin{aligned} \mathbf{\Sigma}_t = E\left[\mathbf{H}^H \mathbf{H}\right] &= \mathbf{U}_t \, E\left[\mathbf{H}_{\mathsf{ind}}^H \mathbf{H}_{\mathsf{ind}}\right] \mathbf{U}_t^H = \mathbf{U}_t \mathbf{\Lambda}_t \mathbf{U}_t^H \\ \mathbf{\Sigma}_r = E\left[\mathbf{H} \mathbf{H}^H\right] &= \mathbf{U}_r \, E\left[\mathbf{H}_{\mathsf{ind}} \mathbf{H}_{\mathsf{ind}}^H\right] \mathbf{U}_r^H = \mathbf{U}_r \, \mathbf{\Lambda}_r \, \mathbf{U}_r^H \end{aligned}$$

where $\mathbf{\Lambda}_t = E\left[\mathbf{H}_{\mathsf{ind}}^H \mathbf{H}_{\mathsf{ind}}\right]$ and $\mathbf{\Lambda}_r = E\left[\mathbf{H}_{\mathsf{ind}} \mathbf{H}_{\mathsf{ind}}^H\right]$ are diagonal. Under certain special cases, the model in (2) reduces to the well-studied i) i.i.d. model, ii) separable correlation model, and iii) virtual representation framework. In particular, the separable model corresponds to $\mathbf{H} = \mathbf{U}_r \, \mathbf{\Lambda}_r^{1/2} \, \mathbf{H}_{\mathsf{iid}} \, \mathbf{\Lambda}_t^{1/2} \, \mathbf{U}_t^H$ where $\mathbf{H}_{\mathsf{iid}}$ is an i.i.d. matrix (see [10] for details).

In this work, we study the coherent case with perfect statistical knowledge at the transmitter. We also assume a simple linear receiver architecture like the linear minimum mean-squared error (MMSE) receiver. The symbol corresponding to the k-th data-stream is recovered by projecting the received signal $\mathbf{y}$ on to the $N_r \times 1$ vector

$$\mathbf{g}_k = \sqrt{\frac{\rho}{M}} \left(\frac{\rho}{M} \mathbf{H} \mathbf{F} \mathbf{F}^H \mathbf{H}^H + \mathbf{I}_{N_r} \right)^{-1} \mathbf{H} \mathbf{f}_k \tag{3}$$

where $\mathbf{f}_k$ is the k-th column of $\mathbf{F}$. That is, the recovered symbol is $\widehat{\mathbf{s}}(k) = \mathbf{g}_k^H \mathbf{y}$. The signal-to-interference-noise ratio (SINR) at the output of the linear filter $\mathbf{g}_k$ is

$$\mathsf{SINR}_k = \frac{1}{\left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1}_{k,k}} - 1. \tag{4}$$

Also, note that the mean-squared error (MSE) of the k-th data-stream, MSE_k, is given by $\left(\mathbf{I}_M + \frac{\rho}{M} \mathbf{F}^H \mathbf{H}^H \mathbf{H} \mathbf{F} \right)^{-1}_{k,k}$.

III. A Case for Structured Precoding

Almost all the current works on precoder design do not assume any specific structure on the precoder matrix $\mathbf{F}$. This is because the main focus is on characterizing the fundamental performance limits of precoding. Although perfect channel information provides a benchmark on the performance, it is difficult to obtain in practice. More importantly, the system performance is not robust under CSI uncertainty. Even a small error in the CSI at the transmitter can lead to a dramatic degradation in performance with a scheme that is designed for the mismatched CSI [8], [9]. Furthermore, even if perfect CSI is available, the complexity in computing the optimal input at the mobile ends may not be low enough to render it implementable [13], [14]. This is because Third Generation wireless systems and beyond are expected to be multi-carrier in nature and the burden of computing the optimal input is magnified by the number of sub-carriers and the rate of evolution of the channel realizations. Besides this, the structure of the input could change, often dramatically, at the rate of evolution of the channel realizations, which also makes it difficult to implement.

The above reasons suggest that a slower rate of adaptation of the input signals, that is of low complexity and is more robust to CSI uncertainty, is preferred. In particular, imperfect channel knowledge arises due to constraints on channel estimation and the quality and feedback of channel and/or statistical feedback. Thus it may not be possible for the family of precoder matrices to have arbitrary structure or allow quantization with arbitrarily fine precision. The additional structure imposed on precoders serves the following purposes: 1) Isolating the impact of inaccuracy in the eigenvectors and the eigenvalues of the precoder on performance with respect to a genie-aided design, 2) Given that there are resource constraints on the reverse link quantization, identifying those features of the channel $\mathbf{H}$ that require an appropriate resource allocation so as to maximize the system performance, and 3) Obtaining more realistic 'intermediate' benchmarks for systems in practice. Much of the focus here is on a specific class of *semiunitary precoder* where $\mathbf{\Lambda_F}$ is fixed to be $\mathbf{I}_M$.

IV. Perfect CSI Structured Precoding

Towards studying the performance of a structured precoding scheme, we first characterize the structure of the optimal signaling scheme when perfect CSI is available at both the transmitter and the receiver. In the unconstrained precoder case, it is well-known that the optimal precoder turns out to be a *channel diagonalizing* structure. That is, the optimal choice of $\mathbf{V_F}$ and $\mathbf{\Lambda_F}$ is such that: $\mathbf{V_F}$ corresponds to the M-dominant eigenvectors in an eigen-decomposition of $\mathbf{H}^H \mathbf{H}$ and the diagonal entries of $\mathbf{\Lambda_F}$ are obtained via waterfilling.

The optimality of the above structure has been proved in [1], [2] with the design metric being the average MSE of the data-streams. Other design metrics where the channel diagonalizing structure is optimal include weighted MSE of the data-streams [3], determinant of the MSE matrix and with a peak-power constraint. In [4], a unified convex programming framework for precoder optimization is proposed by studying two broad classes of functions: Schur-concave and Schur-convex functions. In [4], the authors show that most of the above design criteria can be formulated as either a Schur-concave or Schur-convex function of the MSE and the channel diagonalizing structure is optimal for either case.

When the precoder is structured, it is intuitive (but not obvious) to expect a channel diagonalizing structure to be optimal. The main result of this section is the following theorem.

Theorem 1: Let an SVD of $\mathbf{H}$ be given by $\mathbf{H} = \mathbf{U_H} \mathbf{\Lambda_H} \mathbf{V_H}^H$ with singular values arranged in non-increasing

[3]An $N_t \times M$ matrix $\mathbf{X}$ with $M \leq N_t$ is said to be semiunitary if it satisfies $\mathbf{X}^H \mathbf{X} = \mathbf{I}_M$.

order and $\mathbf{v}_i$ denoting the i-th column of $\mathbf{V_H}$. Let the structure of the precoder be $\mathbf{F} = \mathbf{V_F}\, \mathbf{\Lambda}_{\mathsf{fixed}}^{1/2}$ where $\mathbf{\Lambda}_{\mathsf{fixed}}$ is some fixed matrix of rank M, albeit chosen *arbitrarily*. That is, in the ensuing optimization $\mathbf{\Lambda}_{\mathsf{fixed}}$ is fixed and we only optimize over $\mathbf{V_F}$. The optimal choice of $\mathbf{V_F}$ depends on the nature of the objective functions.

Schur-concave objective functions (and in particular, mutual information with Gaussian inputs) are optimized by

$$\mathbf{V}_{\mathsf{opt}} = \exp(j\theta)\,[\mathbf{v}_1\ \cdots\ \mathbf{v}_M],\ \theta \in \mathbb{R}. \tag{5}$$

Schur-convex objective functions (and in particular, the average uncoded error probability as well as the error probability of the weakest data-stream) are optimized by

$$\mathbf{F}_{\mathsf{opt}} = \exp(j\theta)\,[\mathbf{v}_1\ \cdots\ \mathbf{v}_M]\ \mathbf{\Lambda}_{\mathsf{fixed}}^{1/2}\ \mathbf{\Gamma},\ \theta \in \mathbb{R} \tag{6}$$

for an appropriately chosen unitary matrix $\mathbf{\Gamma}$.

Proof: The results follow by using standard arguments in weak-majorization theory. See [10] for details. ■

Note that the matrix $\mathbf{\Gamma}$ in the definition of $\mathbf{F}_{\mathsf{opt}}$ in (6) can be ignored since $\mathbf{s}$ is i.i.d. and therefore, $\mathbf{\Gamma s}$ is. Thus, we see that in the structured precoding case, the channel diagonalizing structure is still optimal. Towards applying the above result, note that a large class of useful functions satisfy the Schur-concavity property. For example, from [4], we see that any weighted arithmetic or geometric mean of $\{\mathsf{MSE}_k\}$ (with weights chosen appropriately) is Schur-concave. Despite this generality, not all functions are either Schur-concave or Schur-convex. In particular, the mutual information function *cannot* (in general) be expressed as a Schur-concave (or a Schur-convex) function of MSE. Besides mutual information, uncoded error probability is another important metric that describes the performance of a communication system. Our study addresses this metric also. All the above results can be specialized to the case of semiunitary precoders.

V. Asymptotic Performance of Semiunitary Precoding

We assume that instantaneous channel information is not available at the transmitter, but channel statistics are known, and focus on semiunitary precoding. Here, we communicate by precoding along $\mathbf{F}_{\mathsf{semi}}$, a set of M dominant eigenvectors of $\mathbf{\Sigma}_t$ with equal power on the spatial modes. For a given channel realization, let $I_{\mathsf{stat,\,semi}}(\rho)$ and $P_{\mathsf{err,\,stat,\,semi}}(\rho)$ denote the mutual information and error probability achievable with $\mathbf{F}_{\mathsf{semi}}$. Similarly, denote the corresponding quantities with the perfect CSI unconstrained and semiunitary precoders described in Sec. IV and Theorem 1 by $I_{\mathsf{perf,\,unconst}}(\rho)$, $I_{\mathsf{perf,\,semi}}(\rho)$, and $P_{\mathsf{err,\,perf,\,unconst}}(\rho)$, $P_{\mathsf{err,\,perf,\,semi}}(\rho)$, respectively.

It is important to note the distinction between these quantities. While $I_{\mathsf{stat},\,\bullet}(\rho)$ and $P_{\mathsf{err,\,stat},\,\bullet}(\rho)$ are functions of the channel realization $\mathbf{H}$, the precoder structure itself is independent of $\mathbf{H}$, but only dependent on the channel statistics. On the other hand, $I_{\mathsf{perf},\,\bullet}(\rho)$ and $P_{\mathsf{err,\,perf},\,\bullet}(\rho)$ in addition to being dependent on the channel realization also correspond to precoders whose structure is dependent on $\mathbf{H}$ and are chosen optimally.

Average Relative Difference Metrics: The main goal of this paper is to quantify, as a function of the statistics and antenna dimensions,

$$\Delta I_{\mathsf{semi}} \triangleq \frac{E_{\mathbf{H}}\left[I_{\mathsf{perf,\,unconst}}(\rho) - I_{\mathsf{stat,\,semi}}(\rho)\right]}{E_{\mathbf{H}}\left[I_{\mathsf{stat,\,semi}}(\rho)\right]} \tag{7}$$

in the case of mutual information, and

$$\Delta P_{\mathsf{semi}} \triangleq E_{\mathbf{H}}\left[\frac{P_{\mathsf{err,\,stat,\,semi}}(\rho) - P_{\mathsf{err,\,perf,\,unconst}}(\rho)}{P_{\mathsf{err,\,perf,\,unconst}}(\rho)}\right] \tag{8}$$

in the case of error probability. The above metrics correspond to a worst-case measure of relative performance and are more meaningful than a difference metric in studying the relative gap (or closeness) between two signaling schemes, *independent* of the SNR.

Assumptions: The analytical characterization of the two average relative difference metrics in this work critically depends on invoking the assumption of antenna asymptotics. Recent results have shown that the convergence of finite antenna results to asymptotic estimates is usually fast. Thus, our results are meaningful even in the practically relevant small antenna regime. We find it useful to separate our study into two cases: 1) An easily tractable case of *relative receive antenna asymptotics*, where $\frac{M}{N_r} \to 0$, and 2) A more difficult case of *proportional growth of antenna dimensions*, where both $\{M, N_r\} \to \infty$ with $\frac{M}{N_r} \to \gamma$ and $\gamma \in (0,\infty)$ is a constant.

In the error probability case, since exact closed-form expressions are known for the SINR of the data-streams independent of the signaling constellation (see (4)), there is no need to constrain the inputs to be of any particular type. On the other hand, in the case of mutual information, closed-form expressions are not known except for the Gaussian input case. Therefore, we study mutual information with the Gaussian assumption. We also assume that the SNR is reasonably high (more precisely, $\rho \geq \alpha \frac{M}{\mathbf{\Lambda}_t(M)}$ for some suitable $\alpha > 1$) so as to obtain closed-form characterizations of the above metrics. For the sake of simplicity, most of the results will be presented with the separable correlation model. The results can be easily extended to the non-separable case, see [10].

A. Mutual Information Loss

In this section, we focus on the (average) relative loss in mutual information with semiunitary precoding, assuming Gaussian inputs.

Relative receive antenna asymptotics:

Theorem 2: In the separable case,

$$\widetilde{\Delta I}_{\mathsf{semi}} \triangleq E_{\mathbf{H}}\left[\frac{I_{\mathsf{perf,\ semi}}(\rho) - I_{\mathsf{stat,\ semi}}(\rho)}{I_{\mathsf{stat,\ semi}}(\rho)}\right] \leq \frac{2\kappa_1 \sqrt{\sum_{i=1}^{N_r} (\mathbf{\Lambda}_r(i))^2}}{\gamma_r N_r M} \cdot \sum_{k=1}^{M} \frac{1}{\log\left(1 + \frac{\rho}{M}\,\mathbf{\Lambda}_t(k)\right)},$$

where κ_1 is a constant independent of channel statistics. ■

Proportional growth of antenna dimensions:

Theorem 3: In the separable case, let $\{M, N_r\} \to \infty$ with $\frac{M}{N_r} \to \gamma$ and $\gamma \in (0,\infty)$. Also, let the following conditions hold: 1) $\frac{\mathbf{\Lambda}_t(1)}{\mathbf{\Lambda}_t(M)} = \mathcal{O}(1)$, 2) $\frac{\mathbf{\Lambda}_r(1)}{\mathbf{\Lambda}_t(M)} = \mathcal{O}(1)$, 3) $\frac{\mathbf{\Lambda}_r(M)}{\mathbf{\Lambda}_t(M)} = \mathcal{O}(1)$, 4) $\frac{\sum_{k=1}^{M} \mathbf{\Lambda}_t(k)}{\rho_c} = b_1 = \mathcal{O}(1)$, and 5) $\frac{\sum_{k=1}^{M} \mathbf{\Lambda}_r(k)}{\rho_c} = b_2 = \mathcal{O}(1)$. If $\rho \geq \alpha \frac{M}{\mathbf{\Lambda}_t(M)}$ for some $\alpha > 1$, ΔI_{semi} is bounded as

$$\Delta I_{\mathsf{semi}} \leq \frac{\log(e/M) + \kappa_2}{\log(\rho/e) + \frac{1}{M}\sum_{k=1}^{M} \log\left(\frac{\mathbf{\Lambda}_t(k)\mathbf{\Lambda}_r(k)}{\rho_c}\right)} \tag{9}$$

$$\kappa_2 = \kappa_2' + \min\left(E_{\mathbf{H}}\left[\log\left(\frac{\lambda_{\max}(\mathbf{H}_{\mathsf{iid}}^H \mathbf{\Lambda}_r \mathbf{H}_{\mathsf{iid}})}{G_{M,\,\mathbf{\Lambda}_r}}\right)\right],\ E_{\mathbf{H}}\left[\log\left(\frac{\lambda_{\max}(\mathbf{H}_{\mathsf{iid}} \mathbf{\Lambda}_t \mathbf{H}_{\mathsf{iid}}^H)}{G_{M,\,\mathbf{\Lambda}_t}}\right)\right]\right)$$

where κ_2' depends only on the constants in the statement of the theorem, and $G_{M,\,\mathbf{\Lambda}_{\bullet}}$ are the geometric means of eigenvalues, defined as

$$(G_{M,\,\mathbf{\Lambda}_r})^M \triangleq \prod_{k=1}^{M} \mathbf{\Lambda}_r(k), \quad (G_{M,\,\mathbf{\Lambda}_t})^M \triangleq \prod_{k=1}^{M} \mathbf{\Lambda}_t(k).$$

■

It is of interest to understand the structure of the precoding scheme that is optimal from a mutual information viewpoint for a given channel. While many advances have been made along this direction [5], [6], a complete understanding is rendered difficult by the lack of a comprehensive random matrix theory for correlated channels. The above study provides an alternative approach, where we characterize the structure of the channel that is 'best' or 'worst' for a given precoder structure.

Let us freeze $\mathbf{\Lambda}_r$ to be a fixed matrix so as to develop an understanding of the structure of $\mathbf{\Lambda}_t$ that minimizes performance loss. Given that a constraint $\sum_{i=1}^{N_t} \mathbf{\Lambda}_t(i) = \rho_c$ has to be met, it can be checked that performance loss is minimized by the following choice: $\mathbf{\Lambda}_t(1) = \cdots = \mathbf{\Lambda}_t(M) = \frac{\rho_c}{M}$ and $\mathbf{\Lambda}_t(M+1) = \cdots = \mathbf{\Lambda}_t(N_t) = 0$. On the other extreme, the worst choice of $\mathbf{\Lambda}_t$ that maximizes the performance loss is of the form: $\mathbf{\Lambda}_t(1) \approx \rho_c$ and $\mathbf{\Lambda}_t(i) \approx 0, i \geq 2$, but with the added constraint that $\mathsf{rank}(\mathbf{\Lambda}_t) \geq M$. Motivated by Theorem 3, we define a *matching metric for the transmitter side*:

$$\mathcal{M}_t \triangleq \prod_{i=1}^{M} \mathbf{\Lambda}_t(i), \tag{10}$$

that captures the closeness of a given channel from the best and worst channels (characterized above). As $\mathcal{M}_t$ increases, the channel becomes more matched on the transmitter side and the performance loss decreases and *vice versa* (see Fig. 1).

For the impact of $\mathbf{\Lambda}_r$, the above study isolates a *matching metric for the receiver side*, defined as

$$\mathcal{M}_r \triangleq \sum_{i=1}^{N_r} \left(\mathbf{\Lambda}_r(i)\right)^2. \tag{11}$$

Again, with a constraint $\sum_{i=1}^{N_r} \mathbf{\Lambda}_r(i) = \rho_c$ to be met, it can be seen that $\mathcal{M}_r$ is minimized by $\mathbf{\Lambda}_r = \frac{\rho_c}{N_r}\,\mathbf{I}_{N_r}$ and maximized by $\mathbf{\Lambda}_r(1) \approx \rho_c$ and $\mathbf{\Lambda}_r(i) \approx 0, i \geq 2$, but with the added constraint that $\mathsf{rank}(\mathbf{\Lambda}_r) \geq M$. A channel that is matched on both the transmitter and the receiver sides is denoted as a *matched channel* and is optimal for the given precoder structure. The structure of the matched channel can be summarized as: 1) The rank of $\mathbf{\Lambda}_t$ is M with the dominant transmit eigenvalues well-conditioned, and 2) $\mathbf{\Lambda}_r$ is also well-conditioned. A channel that is ill-conditioned on both the transmit and the receive sides such that $\mathsf{rank}(\mathbf{H}) \geq M$ is said to be a *mismatched channel*.

Another interesting consequence is that channel hardening, that occurs as N_r increases, results in the vanishing of ΔI_{semi}. That is, *statistical information is as good as perfect CSI in the receive antenna asymptotics.* This behavior is peculiar of this asymptotic regime [9] and will also be identified in the error probability case.

B. Error Probability Enhancement

We now study the (average) relative error probability enhancement with semiunitary precoding.

Relative Receive Antenna Asymptotics:

Theorem 4: In the separable case, if $\rho \geq \alpha \frac{M}{\mathbf{\Lambda}_t(M)}$ for some $\alpha > 1$, ΔP_{semi} can be bounded as

$$\Delta P_{\mathsf{semi}} \leq \frac{1}{\beta^2 \rho} \cdot \sum_{k=1}^{M} \frac{1}{\mathbf{\Lambda}_t(k)} + \frac{\beta^2 M}{\alpha} + \frac{\beta^2 \rho \sum_{k=1}^{M} \mathbf{\Lambda}_t(k)}{M} \cdot \left(\frac{1}{\alpha} + \frac{1}{\gamma_r}\, \mathcal{O}\left(\frac{\sqrt{N_t} + \sqrt{M}}{\sqrt{N_r}} \right) \right) \tag{12}$$

Thus the dominant term of ΔP_{semi} in the relative antenna asymptotics and large α is of the form: $\frac{1}{\beta^2 \rho} \cdot \sum_{k=1}^{M} \frac{1}{\mathbf{\Lambda}_t(k)} + \beta^2 \frac{\sum_{k=1}^{M} \mathbf{\Lambda}_t(k)}{\mathbf{\Lambda}_t(M)}$. ■

In the non-separable case, we can provide a straightforward extension to ΔP_{semi}.

Theorem 5: Let $\rho \geq \alpha \frac{M}{\gamma_{t,M}} = \alpha \frac{M}{\sum_i \sigma_{iM}^2}$. The dominant term of ΔP_{semi} is bounded as

$$\Delta P_{\mathsf{semi}} \leq \frac{\beta^2}{2} \cdot \frac{\sum_{i=1}^{N_r} \sum_{k=1}^{M} \sigma_{ik}^2}{\sum_{i=1}^{N_r} \sigma_{iM}^2} + \frac{1}{\beta^2 \alpha M} \cdot \sum_i \sigma_{iM}^2 \cdot \sum_{k=1}^{M} \frac{1}{\sum_i \sigma_{ik}^2}. \tag{13}$$

■

As in the mutual information case, we are interested in channels that minimize and maximize the performance loss, ΔP_{semi}. From the above study, it is observed that the choice of $\mathbf{\Lambda}_t$ that minimizes performance loss is such that: 1) It minimizes $\frac{\mathbf{\Lambda}_t(k)}{\mathbf{\Lambda}_t(M)}$, $1 \leq k \leq M$, and 2) It also minimizes $\sum_{k=1}^{M} \frac{1}{\mathbf{\Lambda}_t(k)}$. Both these constraints are met by a channel that maximizes $\mathcal{M}_t$ (as defined in the mutual information case). That is, a channel that is matched on the transmitter side with respect to a mutual information viewpoint is also matched on the transmitter side with respect to an error probability viewpoint. However, it is difficult to make similar conclusions about matching on the receiver side.

On the other hand, note that as the constellation size increases, β decreases. Thus, for any fixed ρ, the first dominant term of ΔP_{semi} in (12) increases as the constellation size increases, whereas the second term decreases. The tension between the two dominant terms determines the optimal choice of constellation to use at a fixed SNR over a given channel. In the extreme case of asymptotically high SNR, the first term vanishes and ΔP_{semi} is minimized with the largest constellation available in the signaling set. The optimality of a larger constellation at high-SNR from an error probability viewpoint is to be intuitively expected. Further, as in the mutual information case, channel hardening results in vanishing ΔP_{semi} as N_r increases. It should also be noted that while ΔP_{semi} vanishes linearly with ρ, ΔI_{semi} vanishes logarithmically in ρ.

VI. Numerical Studies

We present numerical studies that focus on the gap in performance between the perfect CSI and the statistical precoders, as a function of the degree of matching of the channel to the precoder structure. We consider 4×4 channels with $M = 2$, and freeze $\mathbf{U}_t$, $\mathbf{U}_r$ to an arbitrary choice in our study. We also freeze $\mathbf{\Lambda}_r$ to $4\,\mathbf{I}_4$ so as to focus on the impact of matching on the transmitter side. Note that the matching metric, $\mathcal{M}_t = \prod_{k=1}^{M} \mathbf{\Lambda}_t(k)$, takes values in the range $(0, 64]$ in our setting. A family of ∼1700 channels (each characterized uniquely by $\mathbf{\Lambda}_t(k)$, $k = 1, \cdots, 4$) is generated such that $\sum_{k=1}^{N_t} \mathbf{\Lambda}_t(k) = \rho_c = 16$ and $\mathcal{M}_t$ takes values over its range. The channels become more matched (on the transmitter side) to the precoder structure as $\mathcal{M}_t$ increases.

While much of our study in the preceding sections is based on asymptotic random matrix theory, Fig. 1 illustrates that the notion of matched channels developed in this work is useful in characterizing performance, even in practically relevant regimes like 4×4 channels. Fig 1 illustrates that ΔI_{semi} decreases as the channel becomes more matched on the transmitter side for three choices of ρ, whereas Fig 2 illustrates the same trend for ΔP_{semi}. Note that for a given channel as ρ increases, ΔI_{semi} decreases whereas ΔP_{semi} increases. This is because of the contrasting behaviors of $I_{\mathsf{stat,\,semi}}(\rho)$ and $P_{\mathsf{err,\,perf,\,unconst}}(\rho)$ as ρ increases.

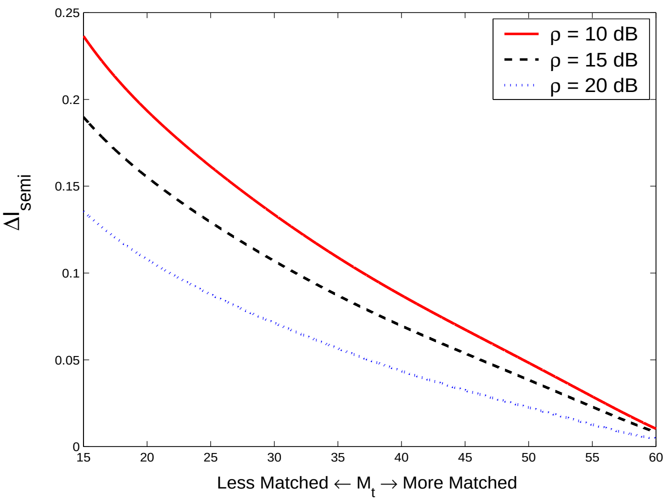

Fig. 1. Gap in average mutual information between statistical and perfect CSI precoding as a function of the matching metric, $\mathcal{M}_t$.

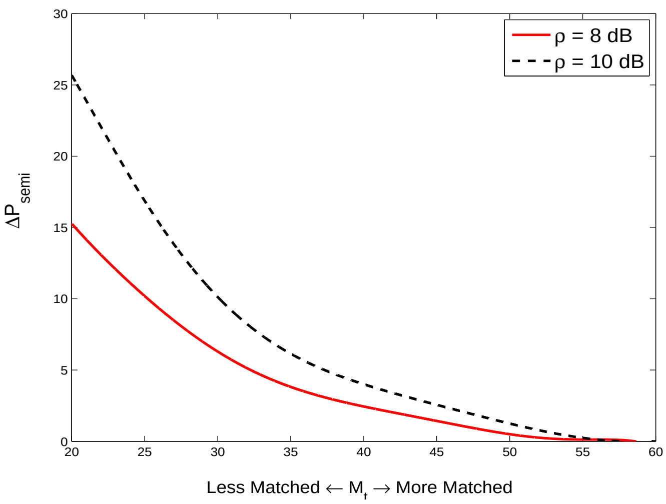

Fig. 2. Gap in uncoded error probability between statistical and perfect CSI precoding as a function of the matching metric, $\mathcal{M}_t$.

VII. Concluding Remarks

In this work, we have proposed the notion of structured precoding. The need for low-complexity and recent advances in partial CSI-based schemes are the motivations behind this work. In studying the performance of structured precoding, we identified matched and mismatched channels, and matching metrics that capture the degree of matching of a channel to the precoder structure *continuously*. In the case of matched channels, the performance of the low-complexity semiunitary scheme is near-optimal. Thus this regime does not warrant the use of any feedback mechanism to improve performance. On the other hand, in mismatched channels the performance gap could be significant; limited feedback schemes are attractive in this regime (see [8] for the construction of low-complexity limited feedback schemes).

Thus our work suggests that in a correlated channel, the optimal number of modes to be excited at a given SNR should be such that the precoder structure is matched to the channel at that SNR. While the notion of source-channel matching is well-known in many contexts in information theory, it is surprising that this natural framework could also capture the performance of multi-antenna communication schemes.

Acknowledgment

This work was partly supported by the NSF under grant #CCF-0049089 through the University of Illinois, and grant #CCF-0431088 through the University of Wisconsin.

References

[1] J. Yang and S. Roy, "On Joint Transmitter and Receiver Optimization for Multiple-Input-Multiple-Output (MIMO) Transmission Systems," *IEEE Trans. Commun.*, vol. 42, no. 12, pp. 3221–3231, Dec. 1994.

[2] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filter-bank Precoders and Equalizers Part I: Unification and Optimal Designs," *IEEE Trans. Sig. Proc.*, vol. 47, no. 7, pp. 1988–2006, July 1999.

[3] H. Sampath, P. Stoica, and A. Paulraj, "Generalized Linear Precoder and Decoder Design for MIMO Channels Using the Weighted MMSE Criterion," *IEEE Trans. Commun.*, vol. 49, no. 12, pp. 2198–2206, Dec. 2001.

[4] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx Beam-forming Design for Multicarrier MIMO Channels: A Unified Framework for Convex Optimization," *IEEE Trans. Sig. Proc.*, vol. 51, no. 9, pp. 2381–2401, Sept. 2003.

[5] A. J. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity Limits of MIMO Channels," *IEEE Journ. Selected Areas in Commun.*, vol. 21, no. 5, pp. 684–702, June 2003.

[6] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO Wireless Channels: Models and Performance Prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[7] E. Jorscwieck and H. Boche, "Optimal Transmission Strategies and Impact of Correlation in Multiantenna Systems with Different Types of Channel State Information," *IEEE Trans. Sig. Proc.*, vol. 52, no. 12, pp. 3440–3453, Dec. 2004.

[8] V. Raghavan, V. V. Veeravalli, and A. M. Sayeed, "Quantized Multimode Precoding in Spatially Correlated Multi-Antenna Channels," *Preprint*, 2008, Available: [Online]. `http://www.ifp.uiuc.edu/∼vasanth`.

[9] V. Raghavan, R. W. Heath, Jr., and A. M. Sayeed, "Systematic Codebook Designs for Quantized Beamforming in Correlated MIMO Channels," *IEEE Journ. Selected Areas in Commun.*, vol. 25, no. 7, pp. 1298–1310, Sept. 2007.

[10] V. Raghavan, A. M. Sayeed, and V. V. Veeravalli, "Low-Complexity Structured Precoding for Spatially Correlated MIMO Channels," *Preprint*, 2008, Available: [Online]. `http://www.ifp.uiuc.edu/∼vasanth`.

[11] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A Stochastic MIMO Channel Model with Joint Correlation of Both Link Ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.

[12] V. Raghavan, J. H. Kotecha, and A. M. Sayeed, "Canonical Statistical Models for Correlated MIMO Fading Channels and Capacity Analysis," *Preprint*, 2007, Available: [Online]. `http://www.ifp.uiuc.edu/∼vasanth`.

[13] J. D. Meindl and J. A. Davis, "*The* Fundamental Limit on Binary Switching Energy for Terascale Integration (TSI)," *IEEE Journ. Solid State Circuits*, vol. 35, no. 10, pp. 1515–1516, Oct. 2000.

[14] B. Razavi, *RF Microelectronics*, Prentice Hall, Upper Saddle River, NJ, 1998.