# Near-Optimal Codebook Constructions for Limited Feedback Beamforming in Correlated MIMO Channels with Few Antennas

Vasanthan Raghavan, Akbar M. Sayeed, Nigel Boston
University of Wisconsin-Madison, Madison, WI, USA
raghavan@cae.wisc.edu, akbar@engr.wisc.edu, boston@engr.wisc.edu

***Abstract*— Transmit beamforming with receive combining is a low-complexity solution that achieves the full diversity afforded by a multi-antenna channel. Building on our recent result which shows that even channel statistics are sufficient to achieve perfect feedback performance (in the limit of antenna dimensions) with beamforming and combining, we propose near-optimal codebook designs for correlated channels with a focus on few antennas at the transmitter and the receiver. In the process, we refine the answer to the question: When are channel statistics sufficient to achieve near perfect feedback performance? We show that the condition number of the transmit and receive covariance matrices hold the key to this question. We partition the transmit and receive covariance spaces into $4$ regions based on well and ill-conditioning of the covariance matrices and show that the number of bits required for near perfect feedback performance is dependent on the condition numbers of these matrices.**

## I. Introduction

Multiple antennas at the transmitter and the receiver provide a mechanism to increase the reliability of signal reception, or rate of information transfer, or a combination of both these aspects. In this paper we focus on achieving the highest level of reliability (full diversity) by employing multiple antennas at both the ends. A simple low-complexity solution towards this goal is transmit beamforming and receive combining. This technique however requires perfect channel state information (CSI) at both the transmitter and the receiver. While perfect CSI at the receiver maybe a reasonable assumption, such knowledge at the transmitter is infeasible in most practical situations. Recent works have therefore focused attention on enhancing the performance of multi-input multi-output (MIMO) systems by exploiting the limited channel knowledge that is usually available at the transmitter [1], [2], [3], [4], [5]. Towards this goal, various solutions have been studied. These solutions range from pure first/second-order statistical feedback to feedback of quantized instantaneous channel information.

In particular, it had been assumed that significant benefits can be achieved when reliable channel information is available at the transmitter than with pure statistical feedback. A recent result of ours [6] shows that in the limit of antenna dimensions, *even* second-order channel statistics are sufficient to achieve the same performance gains as perfect feedback. Building on [6], in this work, we focus on two inter-related problems: 1) What are the factors that influence the rate of convergence to the asymptotics?, and 2) What insights can be acquired from the asymptotics in the context of a practical system deploying few transmit and receive antennas?

To answer the first question, we refine our results in [6] and show that the conditioning of the transmit and receive covariance matrices control the rate of convergence of performance with pure statistical feedback to that of perfect feedback. In the non-asymptotic case of few transmit and receive antennas, we propose a codebook design methodology that exploits the phenomenon of eigenvector hardening (convergence of the dominant right singular vector of the channel to the statistical direction) in correlated MIMO channels. Numerical studies show that the proposed codebook constructions are near-optimal even at low error probabilities for a variety of correlated channel statistics. Furthermore, the number of feedback bits required for a certain level of diversity performance is determined by the level of ill and well-conditioning of the transmit and receive covariance matrix respectively, and we use this criterion to partition the covariance spaces into $4$ regions where limited feedback performance can be clustered.

In this context, our work is closely related to [5]. However, our work differs from [5] in two fundamental aspects: 1) In contrast to a rotation-based Grassmannian design that exploits the separable nature of channel statistics, we propose a systematic codebook design strategy which is explicitly tailored to the conditioning of the transmit and receive covariance matrices, and 2) Our codebook design is *naturally* extendable to more realistic channel models like the canonical statistical model that do not have a separable correlation structure.

## II. System Model

We consider a single user communication system employing transmit beamforming and receive combining, and assume that signalling is done using $N_T$ transmit and $N_R$ receive antennas. The input-output relationship of this system is

$$y = \mathbf{z}^H \mathbf{H} \mathbf{w} \, x + \mathbf{z}^H \mathbf{n} \quad (1)$$

where $\mathbf{H}$ is the $N_R \times N_T$ channel matrix connecting the transmitter and the receiver, $\mathbf{z}$ is the receive combining vector, $\mathbf{w}$ is the transmit beamforming vector, $x$ is the transmitted symbol from a chosen constellation (QPSK, 16-QAM etc.), and $\mathbf{n}$ is the independent noise added at the receiver.

Ideal channel modeling assumes that the entries of $\mathbf{H}$ are independent and identically distributed (i.i.d.) Gaussian random variables. The i.i.d. channel assumption makes the problem studied mathematically tractable, but is unrealistic in applications where either large antenna spacings or a rich scattering environment are not possible. A more realistic channel model is the often-used normalized Kronecker correlation model. Here the channel is given by

$$\begin{aligned} \mathbf{H} &= \frac{1}{\sqrt{\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]}} \; {\mathbf{\Sigma}_R}^{1/2} \; \mathbf{H}_{\mathrm{iid}} \; {\mathbf{\Sigma}_T}^{1/2} \\ &\overset{(a)}{=} \frac{1}{\sqrt{\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]}} \; \mathbf{U}_R \; \Lambda_R^{1/2} \; \mathbf{H}_{\mathrm{iid}} \; \Lambda_T^{1/2} \; \mathbf{U}_T^H \quad (2) \end{aligned}$$

where $\mathbf{\Sigma}_R = \mathbf{E}\left[\mathbf{H}\mathbf{H}^H\right] = \mathbf{U}_R \Lambda_R \mathbf{U}_R^H$, $\mathbf{\Sigma}_T = \mathbf{E}\left[\mathbf{H}^H\mathbf{H}\right] = \mathbf{U}_T \Lambda_T \mathbf{U}_T^H$ correspond to receive and transmit covariance matrices respectively (alongwith their respective eigen decompositions), $\mathbf{H}_{\rm iid}$ is an i.i.d. random matrix as defined above, and (a) follows from the isotropicity of an i.i.d. channel under a unitary transformation. Note that $\mathrm{Tr}\left(\Sigma_T\right) = \mathrm{Tr}\left(\Sigma_R\right) = \mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]$. The above normalization will be useful in obtaining insights into the perturbation theory of dominant singular vector of channel matrices.

Recent channel measurement campaigns have shown that the Kronecker model is a good fit to model the measured data *only* under certain constrained assumptions on the scattering environment [7]. In such situations, a canonical decomposition of the channel along the transmit and receive covariance bases [8] has been shown to be a better fit than the Kronecker model in predicting system metrics like capacity, probability of error etc. The canonical model is given by $\mathbf{H} = \mathbf{U}_R \ \mathbf{H}_{\rm ind} \ \mathbf{U}_T^H$ where $\mathbf{H}_{\rm ind}$ has independent, but not necessarily identical entries, and $\mathbf{U}_R$ and $\mathbf{U}_T$ are unitary matrices. Under the assumption of separable channel statistics, the canonical model reduces to the Kronecker model.

The crux of our results in [6] and this paper hinges on the fact that the entries of the canonical channel matrix has independent entries. However, for ease of illustration and spatial constraints, we assume that the channel is modeled by the Kronecker model in this paper and our results can be easily extended to the more realistic canonical channel model. Insights on design strategies for the canonical model will be drawn wherever appropriate. We further assume that the receiver has a perfect estimate of the channel, while the transmitter knows the channel statistics (that is, $\Sigma_R$ and $\Sigma_T$). This assumption is quite common in existing multi-antenna system designs. The receiver and the transmitter have the knowledge of a channel statistics dependent codebook, $\mathcal{W}$, of transmit beamforming vectors and for every independent channel realization, the receiver feeds back the label of the optimal beamforming vector from $\mathcal{W}$ via a low-rate feedback channel.

## III. Diminishing Returns of Feedback

Initial works on limited feedback studied the codebook design problem for the i.i.d. channel [1], [2], [3]. A natural codebook design criterion in the i.i.d. case was shown in [2] to be the maximization of minimum distance between beamforming vectors. Love et al. [1] showed that this criterion coincides with the objective of maximizing the average received SNR. If the transmitter has perfect CSI, it is easy to see that the optimal beamforming vector is the dominant right singular vector of $\mathbf{H}$ corresponding to the largest singular value, $\mathbf{w}_1$. Given the transmit beamforming direction, $\mathbf{w}_1$, it is also easy to see that the optimal receive combining vector, $\mathbf{z}_1$, is $\frac{\mathbf{H}\mathbf{w}_1}{||\mathbf{H}\mathbf{w}_1||_2}$.

The main focus of this paper is the design of an optimal statistics-dependent codebook, $\mathcal{W}$, so that a suitably defined distortion metric between perfect and limited feedback is minimized. In our prior work [6], we had showed that a suitable distortion metric is $\mathbf{E}\left[1 - \left|\mathbf{w}_{st}^H \mathbf{w}_1\right|^2\right]$ where $\mathbf{w}_{st}$ and $\mathbf{w}_1$ denote the dominant eigenvector of $\Sigma_T$ and the dominant right singular vector of $\mathbf{H}$, respectively.

The fact that the dominant right singular vector of an i.i.d. channel is isotropically distributed can be exploited to show that packing lines in the $N_T$-dimensional complex sphere (also known as Grassmannian line packing) leads to the optimal codebook design [1]. In the case of correlated channels with only receiver side correlation (that is, $\Sigma_T = I$), the Grassmannian line packing solution can be leveraged to obtain efficient codebook constructions [5]. These initial results are crucially dependent on the fact that a right singular vector of a matrix of the form $\mathbf{H} = \Sigma_R^{1/2} \ \mathbf{H}_{\rm iid}$ is isotropically distributed, that is, it is equally likely to point in any direction in the $N_T$-dimensional sphere.

If $\Sigma_T \neq I$, a rotated and normalized Grassmannian codebook design is proposed in [5]. Perturbation theoretic analysis of the $\Sigma_T \neq I$ channel in [6] shows that the isotropicity property (present in $\Sigma_T = I$) is destroyed and there are dominant peaks in the eigen-domain corresponding to the statistical eigen-directions. This result can be summarized as

*Lemma 1:* Let the channel $\mathbf{H}$ be modeled via the Kronecker correlation form as in (2) with $\Sigma_T \neq \mathbf{I}$. Let $\lambda_T^1 \doteq \lambda_{\max}\left(\Sigma_T\right)$ have an algebraic multiplicity 1. Then

$$\lim_{N_R} \mathbf{E}_{\mathbf{H}}\left[\left|\mathbf{w}_{st}^H \mathbf{w}_1\right|\right] = 1. \tag{3}$$

The above lemma says that as long as there is a dominant transmit eigen-direction, *just* increasing $N_R$ hardens the dominant right singular vector of $\mathbf{H}$ to that of the statistical direction, that is, with a very high probability the dominant right singular vector points in the direction of the statistical singular vector. As a consequence, we have

*Proposition 1:* Let $\mathbf{H}$ be as in Lemma 1. Then the average fractional loss in received SNR in the case of pure statistical feedback is at least $\mathcal{O}\left(\frac{\log(N_T N_R)}{N_R^{1/2}}\right)$.

Thus as $N_R$ increases, the gain obtained via a knowledge of perfect CSI cannot better that obtained via a knowledge of *only* the channel statistics. An important point to note is that the above results are asymptotic in the antenna dimensions. We note that the convergence rate in Proposition 1 is the best that has been obtained with current bounding techniques in probability theory. But this does not imply that the rate shown in Proposition 1 is the tightest that can be obtained.

With this comment in mind, we address the following important questions in this paper: Is the convergence rate in Proposition 1 too slow so that in the realm of practical interest (few transmit and receive antennas), the performance gap between pure statistical beamforming and perfect feedback is substantially large? If so, do these asymptotic results offer any guidance in the design of codebooks for the non-asymptotic case? This is the subject of our next section.

## IV. Impact of $\Sigma_T$ and $\Sigma_R$ on Convergence

To address these questions, we first provide a refinement of Proposition 1 that shows the precise influence $\Sigma_T$ and $\Sigma_R$ have on the rate of convergence of performance with pure statistical beamforming to that of perfect feedback.

*Theorem 1:* Arrange the eigenvalues of $\Sigma_T$ and $\Sigma_R$ such that $\lambda_T^1 \geq \lambda_T^2 \geq \cdots \geq \lambda_T^{N_T}$ and $\lambda_R^1 \geq \lambda_R^2 \geq \cdots \geq \lambda_R^{N_R}$.

Define $\gamma_1$, $\gamma_2$ and $\gamma_2^0$ as follows:

$$\gamma_1 \doteq \frac{\lambda_R^1}{\frac{\sum_k \lambda_R^k}{N_R}}, \quad \gamma_2 \doteq 1 - \frac{\lambda_T^2}{\lambda_T^1},$$

$$\gamma_2^0 \doteq \sqrt{\frac{1}{N_T - 1}\sum_{k=2}^{N_T} \frac{x_k}{(1-x_k)^2}}, \quad x_k = \frac{\lambda_T^k}{\lambda_T^1}. \tag{4}$$

Then the fractional loss of received SNR, $\Delta L$, is given by

$$\Delta L \leq \left(K\gamma_1\gamma_2^0\right)\cdot\frac{\log(N_R)}{N_R^{1/2}} \leq K\,\frac{\gamma_1}{\gamma_2}\cdot\frac{\log(N_R)}{N_R^{1/2}} \tag{5}$$

where $K > 0$ is independent of $\Sigma_T$, $\Sigma_R$, $N_T$ and $N_R$.

*Proof:* The proof proceeds along analogous lines as the proof of Proposition 1, the main statement of [6]. A more careful analysis of the constants is needed to show (5). ∎

Note that a large value of $\gamma_1$ implies that the largest eigenvalue of $\Sigma_R$ is well separated from the average eigenvalue of $\Sigma_R$, that is, $\Sigma_R$ is ill-conditioned. Similarly, a large value of $\gamma_2^0$ is equivalent to a large value of $\lambda_T^k$ with respect to $\lambda_T^1$ for all $k$, that is, well-conditioning[1] of $\Sigma_T$. Thus Theorem 1 shows that ill-conditioning of $\Sigma_R$ and well-conditioning of $\Sigma_T$ slows down the rate of eigenvector hardening.

The above conclusion is intuitive. To understand this, we expand the matrix $\mathbf{H}^H\mathbf{H}$ as

$$\mathbf{H}^H\mathbf{H} = \mathbf{U}_T \underbrace{\frac{1}{\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]}\cdot\Lambda_T^{1/2}\,\mathbf{H}_{\mathrm{iid}}^H\,\Lambda_R\,\mathbf{H}_{\mathrm{iid}}\,\Lambda_T^{1/2}}_{=\mathbf{G}}\,\mathbf{U}_T^H$$

$$\mathbf{G}_{ij} = \frac{1}{\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]}\cdot\Lambda_{T_i}^{1/2}\Lambda_{T_j}^{1/2}\sum_k \Lambda_{R_k}\,\mathbf{H}_{\mathrm{iid}_{kj}}\,\mathbf{H}_{\mathrm{iid}_{ki}}^{\star}$$

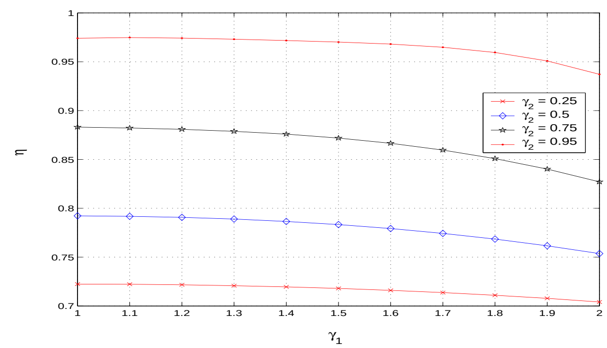
Note that $\mathbf{G}$ is a stochastic diagonally dominant matrix, that is, with a high probability the diagonal entries dominate the off-diagonal entries. This can also be seen from the fact that $\mathbf{E}\left[\mathbf{G}\right] = \Lambda_T$. If $\Sigma_T$ is ill-conditioned, the probability that this stochastic diagonal dominance leads to a dominant eigen-direction along the statistical peak is enhanced. On the other hand, well-conditioning of $\Sigma_T$ implies that more than one statistical direction is dominant and thus beamforming along any fixed direction cannot lead to optimal performance unless we consider the limit of antenna dimensions with the algebraic multiplicity of the dominant direction being 1, the precise conditions of Proposition 1 (See [6] for a detailed discussion).

Furthermore if $i \neq j$, it is easy to see that $\sigma^2 \doteq \left(\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]\right)^2\mathbf{E}\left[\left|\mathbf{G}_{ij}\right|^2\right]$ is given by

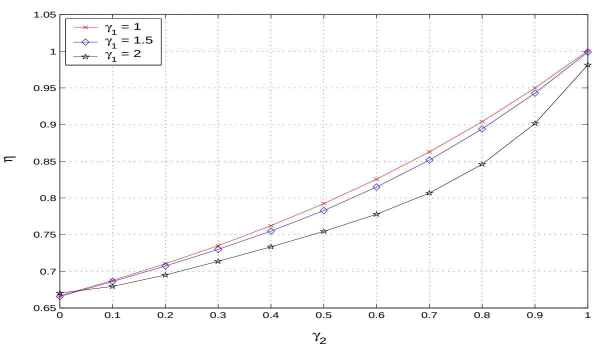
$$\sigma^2 = \Lambda_{T_i}\,\Lambda_{T_j}\sum_{k_1,\;k_2}\Lambda_{R_{k_1}}\,\Lambda_{R_{k_2}}\cdot$$
$$\mathbf{E}\left[\mathbf{H}_{\mathrm{iid}_{k_1,i}}^{\star}\,\mathbf{H}_{\mathrm{iid}_{k_2,i}}\,\mathbf{H}_{\mathrm{iid}_{k_2,j}}^{\star}\,\mathbf{H}_{\mathrm{iid}_{k_1,j}}\right]$$
$$\overset{(b)}{=} \Lambda_{T_i}\,\Lambda_{T_j}\sum_k \Lambda_{R_k}^2 \tag{6}$$

[1] The constant $\gamma_2^0$ is dependent on all the $\lambda_T^k$ and hence precludes insights into limited feedback beamforming. To counter this, we propose a simplified constant $1/\gamma_2$ which is a measure of the well-conditioning of $\Sigma_T$ with a hold on $\lambda_T^1$ and $\lambda_T^2$ alone.

where in (b) we have used the fact that $i \neq j$, and the entries of $\mathbf{H}_{\mathrm{iid}}$ are i.i.d. Now, note that since $\sum_k \Lambda_{R_k} = \mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]$, the variances of the off-diagonal entries is minimized by $\Lambda_R = \frac{\mathbf{E}\left[\mathrm{Tr}\left(\mathbf{H}\mathbf{H}^H\right)\right]}{N_R} I_{N_R}$, or in other words, well-conditioning of $\Sigma_R$ ensures that the perturbations on the dominant right singular vector of $\mathbf{H}$ (from the statistical direction) due to the off-diagonal entries are minimized in the mean-squared sense.




(a) $\eta$ vs. $\gamma_1$




(b) $\eta$ vs. $\gamma_2$

Fig. 1. (a) The plot shows $\eta \doteq \mathbf{E}\;\;\mathbf{w}_{st}^H\mathbf{w}_1$ as a function of $\gamma_1$ for $N_T = N_R = 2$. (b) Here $\eta$ is plotted as a function of $\gamma_2$.

This fact is illustrated in Fig. 1 where $\eta \doteq \mathbf{E}_{\mathbf{H}}\left[\left|\mathbf{w}_{st}^H\mathbf{w}_1\right|\right]$ is plotted as a function of $\gamma_1$ and $\gamma_2$ for $N_T = N_R = 2$. The figure shows that as $\gamma_1, \gamma_2 \rightarrow 1$, corresponding to well and ill-conditioning of $\Sigma_R$ and $\Sigma_T$ respectively, $\mathbf{E}_{\mathbf{H}}\left[\left|\mathbf{w}_{st}^H\mathbf{w}_1\right|\right]$ converges to its maximum value. Thus **pure statistical feedback is sufficient under two cases: 1) $N_R$ is large, or 2) $N_R$ is small with $\Sigma_R$ and $\Sigma_T$ being well and ill-conditioned respectively.** When $N_R$ is small and when such conditioning fails, codebooks have to be designed to achieve better performance gains. From Fig. 1, we can also see that the conditioning of $\Sigma_T$ has a far more significant impact on eigenvector hardening than that of $\Sigma_R$.

### *Heuristic Behind Codebook Construction*

In this section, our focus is on cases not addressed in the previous section: either $\gamma_1$ is sufficiently large, or $\gamma_2$ is sufficiently small with $N_T$ and $N_R$ being small. For other cases, even statistical beamforming is sufficient for near-optimal performance. Theorem 1 and Fig. 1 offer us some insights into the design of optimal codebooks for transmit beamforming in this case. In these cases, even though $\mathbf{E}\left[\left|\mathbf{w}_{st}^H\mathbf{w}_1\right|\right]$ may not be close to 1, this quantity is usually sufficiently large. What this means is that with a high probability, the dominant right singular vector of $\mathbf{H}$ points in the direction corresponding to that given by the channel statistics.

The discrepancy of this average from $1$ is the occurrence of two "rare" events: 1) there are channel realizations where $\mathbf{w}_1$ points locally around $\mathbf{w}_{st}$ but not in the precise direction, and 2) there are some channel realizations, albeit relatively rare, where the angle between $\mathbf{w}_1$ and $\mathbf{w}_{st}$ is large. The probability that these two events occur is small and this probability converges to $0$ in the asymptotics of $N_T, N_R$ as described in [6]. It is also important to note that reliable data reception at high SNR is governed by these rare events. Any codebook design optimized to minimize the probability of error should account for these distortion-inducing channel realizations. We now propose a codebook design for correlated channels based on these principles and we show by numerical studies that the codebook design thus proposed achieves near-optimal performance.

## V. Codebook Design for Correlated Channels

### A. Codebook Construction

The construction elucidated here leads to a codebook of $M_D + M_L + M_G$ code-words. The indices $D, L$ and $G$ stand for dominant, local and global respectively and will be explained in the following three steps of codebook construction. To keep notations simple, we will represent $\lambda_T^i$ by $\mathbf{c}_i$.

Step 1 - **Dominant Statistical Directions**: In the initial step, the codebook is populated with the dominant eigenvectors of $\Sigma_T$. One simple method of implementing this is to choose the eigenvectors corresponding to those eigenvalues which satisfy $\frac{\lambda_T^i}{\lambda_T^1} > \chi\left(\gamma_1, \gamma_2^0\right)$ where $\chi(\cdot)$ is a thresholding function determined by the channel conditioning parameters $\gamma_1$ and $\gamma_2^0$. Smaller the $\chi(\cdot)$ value, larger is $M_D$ and vice versa. For e.g., if $\Sigma_T$ is well-conditioned, for any fixed $\chi(\cdot)$ value, a large fraction of the eigenvalues would satisfy the above condition. In this case, the optimal $\chi(\cdot)$ to be used will depend on the trade-off between a larger codebook and the gain in probability of error performance.

Step 2 - **Local Perturbations**: In this step, we pick $M_{L_i}$ vectors $(i = 1 \ \cdots \ M_D)$ around each dominant statistical direction, $\mathbf{c}_i$, to account for those channel realizations that steer the dominant singular vector in a local neighborhood of $\mathbf{c}_i$. We will identify these local codevectors with $\mathbf{v}_1^i \ \cdots \ \mathbf{v}_{M_{L_i}}^i$. Note that $M_{L_i}$ are non-increasing in $i$ (since the less dominant an eigenvector, the smaller the level of local perturbations around it that is relevant). We define $M_L \doteq \sum_{i=1}^{M_D} M_{L_i}$.

We first pick three parameters: $\epsilon_0$ and $\theta_0$ that determine the perturbation and locality of the $j$-th local codebook candidate $\mathbf{v}_j^i$ about $\mathbf{c}_i$, and $\epsilon$ that ensures that the vectors thus defined are well-separated. The iterative process that leads to the local codebook is as follows: Pick a random vector $\mathbf{r}_1$ such that $\mathbf{v}_1^i \doteq \frac{\mathbf{c}_i + \epsilon_0 \ \mathbf{r}_1}{\|\mathbf{c}_i + \epsilon_0 \ \mathbf{r}_1\|_2}$, $\left|\mathbf{c}_i^H \mathbf{v}_1^i\right| \geq \cos(\theta_0)$, that is, $\mathbf{v}_1^i$ is a unit-normed vector within the cone around $\mathbf{c}_i$ determined by $\theta_0$, as illustrated in Fig. 2. Given $\mathbf{v}_1^i \ \cdots \ \mathbf{v}_j^i, \ \ j < M_{L_i}$, we now describe how to obtain $\mathbf{v}_{j+1}^i$. Define the set $S_{j+1}^i = \bigcup_{k=1}^{j} \mathbf{v}_k^i$. A random vector $\mathbf{r}_{j+1}$ is chosen to define candidate $\mathbf{v}_{j+1}^i$ as $\mathbf{v}_{j+1}^i \doteq \frac{\mathbf{c}_i + \epsilon_0 \ \mathbf{r}_{j+1}}{\|\mathbf{c}_i + \epsilon_0 \ \mathbf{r}_{j+1}\|_2}$. This random vector $\mathbf{r}_{j+1}$ has to be such that

$$\left|\mathbf{c}_i^H \mathbf{v}_{j+1}^i\right| \geq \cos(\theta_0) \quad (7)$$

$$\min_{\mathbf{x} \in S_{j+1}^i} \left|\mathbf{x}^H \mathbf{v}_{j+1}^i\right| \leq \epsilon. \quad (8)$$

Condition (7) leads to localizing $\mathbf{v}_{j+1}^i$ within the cone as in Fig. 2 while Condition (8) leads to a prescribed separation (of at least $\cos^{-1}(\epsilon)$) between the local vectors already designed. Alternately, Condition (8) can be implemented as $\min_{\mathbf{x} \in S_{j+1}^i} \|\mathbf{v}_{j+1}^i - \mathbf{x}\|_2 \geq \epsilon_1$ for a suitable $\epsilon_1$. The algorithm implemented in this paper uses this alternative approach. By optimizing over the three parameters $\epsilon_0, \epsilon_1$ and $\theta_0$, an optimal local set of codevectors can be obtained. The local codebook selection process can also be viewed as a localized Grassmannian line packing solution.

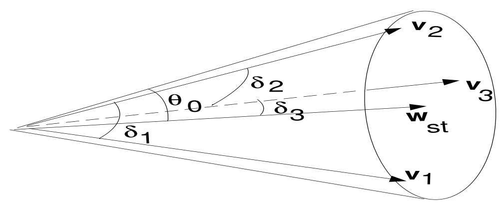



Fig. 2. Local code-word selection algorithm. Given an $\epsilon_0, \theta_0$ and $\epsilon_1$, the codevectors $\mathbf{v}_j^i$ are chosen such that the angle between the various vectors, $\delta_k$, (only a sample of which is represented here) are such that $\delta_k \geq \cos^{-1}(\epsilon_1)$.

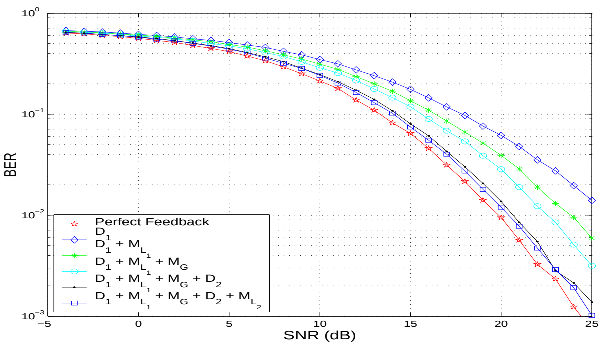
Step 3 - **Global Perturbations**: We now define a set of $M_G$ codevectors that account for those channel realizations that lead to a large perturbation of the dominant right singular vector from the statistical direction. These vectors are chosen so as to maximize the minimum distance between themselves and that of the previously designed code-words. In particular, these vectors could be chosen via the well-developed Grassmannian line packing construction [1] or via random vector quantization [3]. In the few antennas case, Grassmannian packing leads to the optimal choice for the global codevectors.

### B. Numerical Studies

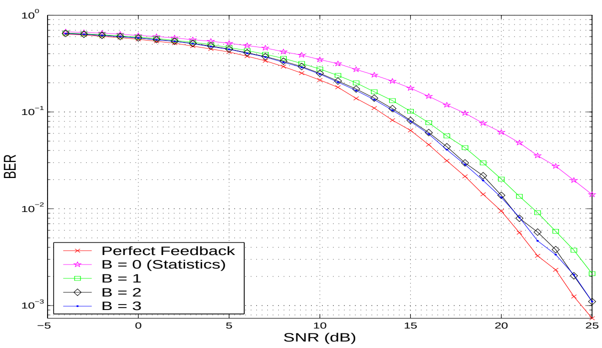
We illustrate the performance of our codebook construction with a representative case. We consider a $2 \times 2$ channel given by the Kronecker model in (2) with conditioning[2] parameters $\gamma_1 = 1.92$ and $\gamma_2 = 0.15$. The codebook design parameters used are: $\epsilon_0 = 0.1, \cos(\theta_0) = 0.92$ and $\epsilon_1 = 0.2$. Signalling is done by using a QPSK constellation. A codebook with 10 codevectors ($M_D = 2, M_{L_1} = M_{L_2} = 3, M_G = 2$) is designed via the procedure explained in the previous section. Fig. 3(a) shows the BER curves with increasing codebook size. In particular, the performance enhancement due to addition of the following components (sequentially) in shown in the figure: 1) The dominant eigenvector ($D_1$), 2) The local codevectors around $D_1$, 3) the global codevectors, 4) the second eigenvector ($D_2$), and 5) the local codevectors around $D_2$.

[2]Note that the maximum and minimum values of $\gamma_1$ and $\gamma_2$ are 2 and 0 respectively. The example considered for numerical study is a particularly pathological example where the rate of eigenvector hardening is minimized.

As can be seen from the figure, optimal signalling with only channel statistics (beamforming along $D_1$) is at least 6 dB away from perfect feedback at $10^{-3}$ probability of error. Following the discussion in Section IV, this is understandable since $N_T = N_R = 2$ and $\frac{\gamma_1}{\gamma_2} \approx 13$. However with 3 bits of feedback in the sequential design, we are able to approach within a fraction of a dB of perfect feedback at the same probability of error.




(a) Performance with Codebook Construction




(b) Optimal Codebook Design

Fig. 3. (a) The figure illustrates the performance of the proposed algorithm in a $2 \times 2$ channel with $\frac{\gamma_1}{\gamma_2} \approx 13$. (b) This figure shows the BER with optimal codebook designs for a fixed feedback constraint on the reverse link.

In multi-user communication systems, even 3 bits of feedback may be prohibitively expensive. Motivated by this concern, we studied the problem: What is the optimal codebook design for $B$ bits of feedback? The results are elucidated in Fig. 3(b). As can be seen, with 1 bit of feedback, we can approach within a couple of dB of perfect feedback. The optimal codebook for this case is achieved by picking one of $D_1$ or $D_2$. This is intuitive since $\frac{\lambda_T^2}{\lambda_T^1} \approx 0.85$. With 2 bits, however, we can near perfect feedback within a dB. In this case, the optimal codebook consists of $D_1$, $D_2$, a codebook vector from the local set of $D_1$, and a codevector from the global set. This optimal codebook design illustrates the heuristic behind the proposed sequential construction (Section V). While the peaks of the probability density function of $\mathbf{w}_1$ occur along the statistical directions, the local and global perturbations are best quantized by the local and global codevectors.

## VI. Discussion

Partitions of Covariance Spaces: The discussion in Section IV shows that the transmit and receive covariance spaces can be partitioned into 4 regions based on the conditioning of these matrices: 1) $\Sigma_R$ is well and $\Sigma_T$ is ill-conditioned, 2) both $\Sigma_R$ and $\Sigma_T$ are ill-conditioned, 3) both $\Sigma_R$ and $\Sigma_T$ are well-conditioned, and 4) $\Sigma_R$ is ill and $\Sigma_T$ is well-conditioned. Note that in Case 1), even statistical feedback is sufficient whereas in Cases 2) and 3), the amount of feedback necessary to achieve within a particular fraction of perfect feedback is non-negligible. This feedback requirement is maximized in Case 4). We conjecture that the amount of feedback required to achieve a particular distortion (for an appropriate distortion metric) increases as $\Sigma_R$ and $\Sigma_T$ become more ill and well-conditioned respectively. Also note that the numerical studies performed in this paper are constrained to Case 4), which has the worst-case feedback requirement.

Implications for the Canonical Statistical Model: The canonical statistical model is a generalization of the Kronecker model and a better fit for measured channel data [8], [7]. We can now leverage the insights obtained for the Kronecker model as to when is statistical feedback sufficient for the case of the canonical model.

*Theorem 2:* For a channel $\mathbf{H}$ modeled by the canonical model, note that the eigenvalues of the transmit and receive covariance matrices are given by $\lambda_{R_i} = \sum_j \sigma_{ij}^2$, and $\lambda_{T_j} = \sum_i \sigma_{ij}^2$. Let $\Lambda_R = \text{diag}(\lambda_{R_i})$ and $\Lambda_T = \text{diag}(\lambda_{T_i})$. Then, pure statistical feedback is sufficient to achieve perfect channel information performance under two conditions: 1) $N_R$ is large, and 2) If $N_R$ is small, and both a) $\Lambda_T$ is ill conditioned and b) $\Lambda_R$ is well conditioned.

*Proof:* The proof is skipped due to spatial constraints. The intuition behind the proof is easy to describe. Condition 1) follows from [6] while Condition 2a) ensures that with a high probability the matrix $\mathbf{H}_{\text{ind}}\mathbf{H}_{\text{ind}}^H$ is diagonally dominant and Condition 2b) implies that the variances of the off-diagonal entries of $\mathbf{H}_{\text{ind}}\mathbf{H}_{\text{ind}}^H$ are minimized, thus constraining the perturbations stochastically. ■

When either of these two conditions 2a) and 2b) are not met, and when $N_R$ is small, codebooks have to be designed to near the performance with perfect channel knowledge. Codebook design methodologies follow analogous to the discussion in Section V. Numerical studies are not reported here due to spatial constraints.

## References


[1] D. J. Love, R. W. Heath Jr., T. Strohmer, "Grassmannian Beamforming for Multiple-Input Multiple-Output Wireless Systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2735-2747, Oct. 2003

[2] K. K. Mukkavilli, A. Sabharwal, E. Erkip, B. Aazhang, "On beamforming with finite rate feedback in multiple-antenna systems," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2562-2579, Oct. 2003

[3] W. Santipach, Y. Sun, M. Honig, "Benefits of Limited Feedback for Wireless Channels," *Proc. 2003 Allerton Conference*, Oct. 2003

[4] J. C. Roh, B. D. Rao, "Vector Quantization Techniques for Multiple-Antenna Channel Information Feedback," *SPCOM*, Bangalore, India, Dec. 2004

[5] D. J. Love, R. W. Heath Jr., "Grassmannian Beamforming on Correlated MIMO Channels," *Proc. Globecom*, Dallas, 2004.

[6] V. Raghavan, A. M. Sayeed, N. Boston, "When is Limited Feedback for Transmit Beamforming Beneficial?" *Proc. ISIT* , 2005.

[7] H. Ozcelik, M. Herdin, W. Weichselberger, J. Wallace, E. Bonek, "Deficiencies of 'Kronecker' MIMO Radio Channel Model," *Elec. Lett.*, vol. 39, no. 16, pp. 1209-1210, Aug. 2003

[8] J. H. Kotecha, V. Raghavan, A. M. Sayeed, "Canonical Statistical Modeling and Capacity Analysis of Correlated MIMO Fading Channels," *Submitted to the IEEE Trans. Wireless Comm.*, Mar. 2005