When is Limited Feedback for Transmit Beamforming Beneficial?

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 and receive combining are low complexity techniques that help in achieving the full diversity afforded by the multi-antenna channel. Complete channel information at the transmitter may be possible only under rare instances. Thus quantized beamforming with limited feedback on a reverse link has been a topic that has attracted great attention recently. But almost all of the work to date has focused on modeling the channel with independent, identically distributed (i.i.d.) Rayleigh fading between antenna pairs. In this paper, we consider the correlated channel case, and show that Grassmannian line packing is an artificial artifact of the i.i.d. assumption. We show that there are dominant peaks in the eigen-domain when correlation is imposed and the code-book construction should be matched to the correlation in the channel, which renders the problem in this case easier than for i.i.d. channels.

I. INTRODUCTION

Multi-antenna systems have come to the forefront of research with the seminal works of Telatar [1], and Foschini and Gans [2] who have shown that enormous capacity increase is possible under rich scattering conditions. Exploiting the spatial dimensions to achieve enhanced resilience to fading has been the other extreme objective. The tradeoff between achieving the full rate and the full diversity in a multi-antenna system has been analyzed in [3].

Diversity in a multi-antenna system is achievable either through space-time coding or using the channel state information at the transmitter intelligently. Low complexity techniques like transmit beamforming and receive combining have attracted significant attention of late. This is because of the possibility to feed back at a lower rate partial channel state information from the receiver to the transmitter, both in deployed 3-G wireless systems and in ongoing standardization efforts for 4-G systems.

Recent works towards this goal have addressed the problem of code-book design for the transmit beamforming vector under the independent and identically distributed (i.i.d.) channel assumption [4], [5]. A natural code-book design criterion in the i.i.d. case was shown in [5] to be maximization of minimum distance between beamforming vectors. Love et al. [4] showed that this criterion coincides with the objective of maximizing the average SNR at the receiver.

Most realistic channels are far from having i.i.d. Rayleigh fading between antenna pairs. Therefore existing code-book designs have to be modified to account for channel correlation to achieve higher performance gains in realistic channel conditions. But before pursuing the objective of efficient code-book designs for the correlated channel case, one needs to address the fundamental question: When is limited feedback for transmit beamforming beneficial? This paper concerns this important question.

We show that the i.i.d. channel assumption, far from making the code-book design problem easier, masks it due to the isotropic nature of the channel, i.e. the singular vectors of the channel are equally probable to point in any direction in the eigen-domain. We first consider a Kronecker product correlation model, and later an even more realistic virtual representation of channel matrices to show that correlation in the channel entries renders the code-book design problem easy. Unlike the i.i.d. channel which has singular vectors isotropically distributed, a correlated channel has peaks in the eigen-domain corresponding to the eigenvectors of the transmit covariance matrix. This observation leads to an important, albeit previously unnoticed, conclusion: Limited feedback for transmit beamforming is beneficial when the transmit covariance matrix is the identity matrix [6] (with the i.i.d. channel being a special case of this condition), and in other cases provided the number of receive antennas are few.

The system model under consideration is described in the next section, while mathematical preliminaries are introduced in Section III. Code-book design criterion and conditions under which limited feedback is beneficial are considered in Section IV, while conclusions are drawn in Section V. Some of the proofs have been relegated to the appendix. In this paper we use the following notations: $x_i$ to denote the $i$-th entry of the vector $x$, $A_{ij}$ for the $i,j$-th entry of the matrix $A$, $E[A]$ for the expectation operator with respect to the random variable $A$, $|| \cdot ||_2$ to refer to the 2-norm of a vector or a matrix, $H$, $T$ for Hermitian transpose, and regular transpose respectively.

We also remind the readers that we mean $f(N) = O(g(N))$ by the notation $f(N) = O(g(N))$.

II. SYSTEM MODEL

A. System Setup

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

$$ y = z^H H w x + z^H n $$

(1)

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

We assume that the receiver knows the channel perfectly, while the transmitter knows the channel statistics. This assumption is quite common in existing multi-antenna system designs. The receiver and the transmitter have the knowledge of a finite code-book, $W$, of transmit beamforming vectors.
(dependent on the channel statistics) and the receiver feeds back for every independent channel realization, the optimal beamforming vector from this code-book via a low-rate feedback channel (viz. the vector index in the code-book). We will follow the terminologies of transmit beamforming as in [4].

The beamforming and combining vectors are chosen to maximize the received SNR, $\gamma$, which is,

$$\gamma = \frac{E_s |z^H H w|^2}{||z||^2 N_0}$$  \hspace{1cm} (2)

where $E_s = \mathbb{E}|x|^2$ is the symbol energy, and $N_0$ is the noise variance. If the transmitter has CSI, then it is easy to see that the optimal beamforming vector is the dominant right singular vector of $H$ corresponding to the largest singular value. Given a transmit beamforming vector, it is also easy to see that the optimal receive combining vector is $\frac{H w}{||H w||_2}$. The distortion, $G(W)$, between the system with complete CSI and the system employing limited feedback is given by [4]

$$G(W) = \mathbb{E}_H [\lambda_1 - ||H Q_w(H)||^2_2]$$  \hspace{1cm} (3)

where $\lambda_1$ is the largest eigenvalue of $H^H H$, and $Q_w(H) = \arg \max_{w \in W} ||H w||^2_2$. \hspace{1cm} (4)

Here $\lambda_1$ is the normalized gain achieved by a system that has complete CSI and $||H Q_w(H)||^2_2$ is the gain corresponding to limited feedback. The design objective thus is to select a code-book based on the channel statistics so as to minimize the distortion measure.

**B. Channel Models**

In this sub-section we consider various statistical and parametric modeling schemes for the channel $H$. The statistics of $H$ depend on antenna geometry, physical scattering environment, frequency of operation etc. Ideal channel modeling assumes that the entries of $H$ are i.i.d. Gaussian random variables [1], [2]. Analysis of an i.i.d. channel is tractable, but the i.i.d. channel assumption, however, may not be realistic for wireless applications where large antenna spacings or a rich scattering environment are not possible.

Parametric physical models, on the other hand, explicitly model signal copies from different directions [7]. Despite being highly accurate in capturing the scattering features, most of these models suffer from the problem of mathematical tractability. They do not offer insights into designing effective signaling strategies or space-time code construction in the general case. The need for channel modeling schemes that capture most of the statistical information of the underlying physical environment, and yet stay within the realms of mathematical tractability is therefore paramount.

Towards this goal, we first study the often-used Kronecker product (or separable) correlation model. This channel model is characterized by a stacked covariance matrix that has a Kronecker product structure. More precisely, the channel matrix $H$ is of the form

$$H = \Sigma_R^{1/2} H_{iid} \Sigma_T^{1/2}$$  \hspace{1cm} (5)

where $\Sigma_R$ and $\Sigma_T$ correspond to receive and transmit covariance matrices respectively, and $H_{iid}$ is an i.i.d. random matrix as defined earlier. If $h = \text{vec}(H)$ is the vector obtained by stacking the columns of $H$, then it follows that

$$\mathbb{E} [hh^H] = \Sigma_T \otimes \Sigma_R.$$  \hspace{1cm} (6)

The Kronecker product correlation model, despite its mathematical tractability and simplification of performance analysis of MIMO systems, suffers from deficiencies acquired by the separability property of channel correlation that limits the number of degrees of freedom in the model [8], [9]. Measurement campaigns have shown that the Kronecker product correlation model is accurate in capturing the underlying channel statistics under certain conditions [10], but in general, the model underestimates the measured channel capacity [8] and other system parameters.

Various statistical models have been proposed to increase the number of degrees of freedom in the channel model. When the transmitter and receiver antennas form a uniform linear array (ULA), the virtual channel representation [11] has been shown to capture the channel characteristics quite accurately. The virtual representation exploits the finite dimensionality of the signal space, afforded by finite number of antenna elements and finite array aperture, to develop a linear channel representation that uses spatial beams in fixed, virtual directions. The virtual channel representation can be expressed as

$$H = \sum_{m=1}^{N_R} \sum_{n=1}^{N_T} H_{V,m,n} a_R(\theta_{R,m}) a_T^H(\theta_{T,n}) = \tilde{A}_R H_V \tilde{A}_T^H$$  \hspace{1cm} (7)

where $\{\theta_{R,m} = \frac{m}{N_R}\}$ and $\{\theta_{T,n} = \frac{n}{N_T}\}$ are fixed virtual angles that result in unitary $A_R$ and $\tilde{A}_T$. Here, $A_R = [a_R(\theta_{R,1}) \cdots a_R(\theta_{R,N_R})]$, $\tilde{A}_T = [a_T(\theta_{T,1}) \cdots a_T(\theta_{T,N_T})]$, and $H_V$ is the virtual channel matrix.

The virtual channel coefficients represent the coupling between the transmit and the receive virtual angles. The most important advantage of the virtual representation is the fact that the virtual entries are approximately uncorrelated and this uncorrelatedness of channel coefficients gets better as number of antennas increase. The number of virtual channel entries with non-zero variance corresponds to the number of independent degrees of freedom afforded by the channel. The readers are referred to [11] for a detailed treatment of the virtual channel representation in the narrowband case and [15], [16] for the wideband case.

The generalization of the virtual representation to the non-ULA case, the canonical statistical model [12], [13], [14], has also been studied extensively of late, both mathematically and via measurement campaigns. This model assumes that the auto and cross-correlation matrices on both transmitter and receiver sides have the same eigen-basis, and exploits this redundancy to decompose the channel as

$$H = U_R^{1/2} H_{\text{ind}} U_T^{1/2}$$  \hspace{1cm} (8)

where $U_R$ and $U_T$ correspond to receive and transmit covariance matrices respectively, and $H_{\text{ind}}$ is a random matrix with independent entries that are zero mean, proper complex Gaussian and variance $\sigma^2_{ij}$, which are not necessarily equal.
It can be shown that when ULAs are used at the transmitter and receiver ends, the Kronecker product model reduces to a special case of the virtual representation [17]. Besides, in the non-ULA setting, the Kronecker product correlation model is a special case of the canonical statistical model. The canonical statistical model not only incorporates the virtual representation and the Kronecker product model as special cases, but it also has been shown to be quite accurate in predicting performance metrics of measured channels. Accuracy of the canonical statistical model in predicting measured channel capacity and other performance metrics is reported in [13, 18].

III. MATHEMATICAL PRELIMinARIES

In this section, we introduce the mathematical results that would be needed in our code-book design for the correlated channel case. The law of large numbers and the central limit theorem play an ubiquitous role in the analysis of communication systems. The first lemma of this section pertains to the rate of convergence of the law of large numbers for a sum of independent (not necessarily identically distributed) random variables [19].

Lemma 1 (Yuen, 1991): Let \( \{X_i\} \) be a sequence of independent zero-mean random variables and let \( \{a_{nk}\} \) be a double array of real numbers. Let \( \|a_n\|_{\beta} = \left(\sum_k |a_{nk}|^\beta\right)^{1/\beta} \). Suppose there exist real numbers \( \alpha > 0, \beta \geq 1 \) such that \( 0 < \|a_n\|_{\min(\beta,2)} < \infty \) for all \( n \), \( \|a_n\|_{\beta} \to 0 \), and there exists \( \alpha \leq T_{\alpha\beta} \leq \infty \) such that \( f_k(t) = E[e^{tX_k}] \leq e^{\alpha t|t|^\beta} \) for all \( k \), and all real \( t \) such that \( |t| \leq T_{\alpha\beta} \). If \( \beta > 1 \), then

\[
P\left(\sum_{k=1}^n a_{nk}X_k \geq \epsilon\right) \leq 2 \exp\left(-\frac{\epsilon \min(T_{\alpha\beta},T)}{|\|a_n\|_{\beta}} + \alpha \left[\min(T_{\alpha\beta},T)\right]^{\beta}\right)
\]

where \( T = \left(\frac{\epsilon}{\alpha |\|a_n\|_{\beta}}\right)^{1/(\beta-1)} \).

Note that the theorem is stated for real random variables \( X_k \) in the paper, and its extension to the complex case is straightforward. The next lemma of this section concerns the eigenvectors of perturbed matrices. The classical Gershgorin theorem [20] states that the eigenvalue of a matrix perturbed by an error matrix differs from the eigenvalue of the unperturbed matrix by the (spectral-)norm of the perturbation. However, the connection between the eigenvectors of perturbed and unperturbed matrices is more subtle. The famous Davis-Kahan sin(\(\theta\)) theorem [20] states that the sine of the angle between eigenspaces of matrices that are close to one another is related to the norm of the difference between the matrices, provided the eigenvalues of the two matrices are well separated. In our setting, the following easier lemma [21] suffices.

Lemma 2 (Mathias, 1994): Let \( \Lambda = \text{diag}(\lambda_i) \) be a positive definite diagonal matrix with distinct eigenvalues. Let \( X \) be a Hermitian matrix and consider \( S(\epsilon) = \Lambda + \epsilon X \). Then for sufficiently small \( \epsilon \), we have \( \lambda_j(\epsilon) = \lambda_j(S(\epsilon)) \) distinct and one can choose \( \hat{u}(\epsilon) \) to be an eigenvector of \( S(\epsilon) \) such that

\[
\hat{u}(\epsilon)_i = 1 + O(\epsilon^2)
\]

\[
|\hat{u}(\epsilon)_i| \leq \frac{|X_{ij}|}{|\lambda_i - \lambda_j|} + O(\epsilon^2), \quad i \neq j.
\]

We note that in the above lemma, only the largest eigenvalue of \( \Lambda \) needs to be distinct to ensure that the eigenvector corresponding to the largest eigenvalue of \( \Lambda + \epsilon X \) has a perturbation of \( \epsilon \) over the eigenvalue of \( \Lambda \). This is an important point that we would exploit in the sequel.

IV. CODE-BOOK DESIGN

A. Distortion Measure

In this sub-section, we propose a mathematically tractable upper bound for the distortion measure in (3) for a general correlated channel \( H \). We remind the readers that this goal is simplified in the case of an i.i.d. channel, since the eigenvalues and the eigenvectors of \( HH^H \) are independent. However, this independence is no longer true in the case of a correlated \( H \). However, we have the following upper bound for the distortion measure.

Proposition 1: The distortion measure, \( G(W) \) is upper bounded by \( U \) where

\[
U = \sqrt{1 - E_H(A)} \left[ E_H(\lambda_1) + \sqrt{2 \text{Var}_H(\lambda_1)} \right]
\]

where \( \lambda_1 \) is the largest eigenvalue of \( HH^H \), \( u_1 \) is the corresponding eigenvector, and \( A \) denotes \( \max_{w \in W} |u_1^H w|^2 \).

Proof: See Appendix A. \( \square \)

The first part of \( U \), which is \( \sqrt{1 - E_H(A)} \) represents the effect of code-book design on the distortion, and \( E_H(\lambda_1) + \sqrt{2 \text{Var}_H(\lambda_1)} \) corresponds to the channel contribution to distortion. Our focus here will be on minimizing the average fractional loss in SNR, \( 1 - E_H(A) \), so as to reduce distortion. We also note that the upper bound in Proposition 1 complements the upper bound on distortion in the i.i.d. channel case, (18) in [4].

B. Conditions Under Which Limited Feedback is Beneficial

In this sub-section, we explore the conditions under which limited feedback for transmit beamforming produces substantial benefits. To be precise, we consider the case when there is no feedback from the receiver to the transmitter and show that the average fractional loss in SNR diminishes with increasing receiver dimensions. It is easy to see that when there is no feedback between the receiver and the transmitter the optimal beamforming vector is the largest eigenvector of \( \Sigma_T \) (or \( U_T \)), provided such a dominant direction exists. We would then like to quantify the average loss in SNR given that our codebook has only one code vector, the largest eigenvector of \( \Sigma_T \). Under this assumption, the part of the upper bound in Proposition 1 due to the code-book design can be written as \( E_H \left[ 1 - |u_1^H w|^2 \right] \). Our main conclusion stems from the following theorem.

Theorem 1: Let the channel \( H \) be modeled via the Kronecker product correlation form (with \( \Sigma_T \neq I \)). Also assume that the largest eigenvalue of \( \Sigma_T \) has an algebraic multiplicity
1. Then the average fractional loss of SNR in the case of no feedback is $O\left(\frac{\log N_T N_R}{N_T^2} \right)$.

Proof: See Appendix B.

The theorem shows that the average fractional loss in SNR diminishes to zero as the receiver dimensions increase. Intuitively, the independence of channel entries leads to channel hardening [22], which in turn results in the concentration of the dominant right singular vector of $\mathbf{H}$ in the direction corresponding to the largest eigenvector of the transmit covariance matrix. The separability assumption of the largest eigenvalue of $\Sigma_T$ makes intuitive sense since the presence of a dominant eigen-direction implies that when the transmitter has only statistical information about the channel, beamforming along that dominant direction would be optimal. It appears that this multiplicity assumption about the largest eigenvalue is superfluous, for if the eigenspace corresponding to the largest eigenvalue had a higher dimensionality than 1, beamforming along any vector in this space would be optimal.

Also the convergence rate proven in Theorem 1 is critically dependent on the convergence rate of Lemma 1. Better bounds may be obtained in the cases under consideration, there is a dominant peak in the eigen-domain of $\Lambda$ in the direction corresponding to the largest eigenvector of the transmit covariance matrix. Exploiting this knowledge, one could construct sub-optimal designs for the correlated channel.

V. CONCLUSION

In this paper, we have studied the problem of limited feedback beamformer design for the multi-antenna channel. In this work, we considered the beamformer design problem in the correlated channel situation, and contrasting the i.i.d. case, we have showed that the loss in performance when no feedback is done vanishes as the number of receiver antennas increase. This result of ours shows that limited feedback offers substantial benefits in two scenarios: the case where $\Sigma_T = \mathbf{I}$ (i.i.d. channel being a special case of this) that has been already considered in the literature [6], and in the general correlated setting provided the antenna dimensions are few.

Theorem 2: Assume that the channel $\mathbf{H}$ is expressed via a virtual representation or the canonical statistical model.

The proof is exactly similar to the proof of Theorem 1. The assumption of well-separability is used to show that with a very high probability the largest eigenvalue of $\Lambda$ (as in Theorem 1) is distinct.

Theorems 1 and 2 show that in a correlated channel, gains obtained by limited feedback diminish as the receiver dimensions increase. This is however not the case in the case when $\Sigma_T = \mathbf{I}$ (with an i.i.d. channel being a special case of this condition) where there are no dominant singular vectors due to the isotropicity of the channel matrix. In the cases under consideration, there is a dominant peak in the eigen-domain and this peak precisely corresponds to the dominant right singular vector of $\mathbf{H}$ when the antenna dimensions increase.

The limited feedback (code-book design) problem belongs to a category of problems where the solution is easier when the channel is correlated than if it is i.i.d. unlike capacity analysis, space-time code design etc. Thus substantial gains via limited feedback could be hoped for only in the small antenna dimensions.

The code-book design in the correlated channel case is more complicated than the i.i.d. design since not much is known about the probability distribution of the eigenvectors of a random non-isotropic matrix. This knowledge is crucial in constructing optimal code-books with constraints on feedback rate. However as shown in Theorem 1, the largest singular vector of $\mathbf{H}$ tends to converge in some sense to the largest eigenvector of the transmit covariance matrix. Exploiting this knowledge, one could construct sub-optimal designs for the correlated channel.

APPENDIX

A. Proof of Proposition 1

Proof: It is not difficult to see that $G(\mathbf{W})$ can be upper bounded by $E_H \left[ \lambda_1 (1 - |u^H_1 q_1 |^2) \right]$ (see (18) in [4]). An easily tractable upper bound to the former is $E_H \left[ \lambda_1 (1 - \max_{w \in \mathbf{W}} |u^H_1 w^2 |) \right]$. Let $A$ denote the random variable $\max_{w \in \mathbf{W}} |u^H_1 w^2 |$. We now employ the definition of the correlation coefficient between two random variables to obtain

$$E_H \left[ \lambda_1 (1 - A) \right]$$

$$= E_H(\lambda_1)E_H[1 - A] + \rho \sqrt{\text{Var}_H(\lambda_1)} \sqrt{\text{Var}_H(1 - A)}$$

$$\leq E_H(\lambda_1)E_H[1 - A] + \sqrt{E_H(A^2) - [E_H(A)]^2} \sqrt{\text{Var}_H(\lambda_1)}$$

$$\leq E_H(\lambda_1) [1 - E_H(A)] + \sqrt{\text{Var}_H(\lambda_1)} \sqrt{1 - [E_H(A)]^2}$$

$$= \sqrt{1 - E_H(A)} \left[ \sqrt{1 - E_H(A)} E_H(\lambda_1) \right]$$

$$+ \sqrt{1 - E_H(A)} \left[ \sqrt{1 + E_H(A)} \sqrt{\text{Var}_H(\lambda_1)} \right]$$

$$\leq \sqrt{1 - E_H(A)} \left[ E_H(\lambda_1) + \sqrt{2 \text{Var}_H(\lambda_1)} \right]$$

where in (10) we have used the fact that $\rho$ is real and $|\rho| \leq 1$, in (11) we have used a trivial upper bound of 1 for $E_H(A^2)$, and in (12) we have used the fact that $E_H(A) \leq 1$. Thus we are done.

B. Proof of Theorem 1

Proof: Let $U_TD_TU_T^H$ and $U_RD_RU_R^H$ denote the eigen-decompositions of $\Sigma_T$ and $\Sigma_R$ respectively. We first note that exploiting the isotropicity property of an
i.i.d. random matrix, \( \mathbf{H}^H \mathbf{H} \) can be written as
\[
\mathbf{H}^H \mathbf{H} = N_R \left( \frac{1}{N_R} \mathbf{H}^H \mathbf{H} \right) = N_R \mathbf{U}_T \mathbf{D}_T^{1/2} \left( \frac{1}{N_R} \mathbf{H}^H \mathbf{R} \mathbf{H} \right) \mathbf{D}_T^{1/2} \mathbf{U}_T^H
\]
= \( \mathbf{Y} + \Delta \mathbf{Y} \), where \( \mathbf{Y} = N_R \mathbf{U}_T \mathbf{D}_T^{1/2} \mathbf{D}_T^{1/2} \mathbf{U}_T^H \), \( \mathbf{D} = \text{diag}\left( \frac{1}{N_R} \mathbf{H}^H \mathbf{R} \mathbf{H} \right) \), \( \Delta \mathbf{Y} = N_R \mathbf{U}_T \mathbf{D}_T^{1/2} \mathbf{O}_T \mathbf{D}_T^{1/2} \mathbf{U}_T^H \), \( \mathbf{O} = \frac{1}{N_R} \mathbf{H}^H \mathbf{D}_R \mathbf{H} - \mathbf{D} \). That is, we decompose \( \mathbf{H}^H \mathbf{H} \) into a part constituted by the dominant diagonal terms, \( \mathbf{D} \), and the off-diagonal components in \( \mathbf{O} \), which vanish in the limit of large \( N_R \) due to the law of large numbers, placing us within the realm of Lemma 2.

Using the hypothesis that the largest eigenvalue of \( \Sigma_T \) is well separated from the other eigenvalues, and Lemma 1 (with \( \beta = 2 \), \( f_k(t) = \frac{1}{1-t(D_k)} \) [23], \( T_{\alpha, \beta} = \frac{1}{\sqrt{\max_k(D_k)}} \), \( ||a_n||_\beta = \sqrt{N_R} \), and \( \alpha \) sufficiently large), we have that the largest eigenvalue of \( \Sigma_T \mathbf{D} \) is distinct with probability greater than \( p_0 = 1 - 2 \exp\left( -\frac{\gamma^2}{8 \max_k(D_k)} \right) \). Further analysis is done modulo this high probability set. We use \( \Lambda \) to denote \( \Sigma_T \mathbf{D} \), and \( X = \max_k(\mathbf{D}_k) \). Thus \( \mathbf{H}^H \mathbf{H} = N_R \mathbf{U}_T [\Lambda + \max_k \mathbf{O}_k] \mathbf{X} \mathbf{U}_T^H \). Applying Lemma 2, we have the following conclusion:
\[
\hat{u}(k) = 1 + c_k, \ |c_k| = O(\epsilon^2)
\]
\[
\hat{u}(l) = e, \ |e| \leq \epsilon K + O(\epsilon^2), \ l \neq k
\]
with probability \( p = P(\max_k |\mathbf{O}_k| \leq \epsilon) \), where \( \hat{u} \) is the largest eigenvector of \( \Lambda + \max_k \mathbf{O}_k \), \( X \) is the index corresponding to the largest diagonal entry of \( \Lambda \), and \( K \) is a positive constant.

Note that the fractional loss in SNR corresponds to the squared value of the sine of the angle between the eigenvectors of the perturbed and unperturbed matrices. Elementary algebra using the orthogonality of columns of \( \mathbf{U}_T \) shows that the fractional loss in SNR random variable is \( O(\epsilon^2) \) with probability \( \epsilon \). Thus the average fractional loss in SNR is upper bounded by \( \underline{U}_1 = 1 + (K_1 \epsilon^2 - 1)(p + p_0 - 1) \) where \( K_1 \) is a positive constant. Note that \( q = 1 - p \leq \sum_k P(|\mathbf{O}_k| > \epsilon) \leq \frac{N_r(N_r-1) - 1}{2} \). We observe that a product of zero-mean Gaussian random variables (of variances \( \sigma_1^2 \) and \( \sigma_2^2 \), respectively) has a Bessel-function of the zeroth order like probability density function and its Moment Generating Function is given by \( \left( \frac{1}{1-\sigma^2 \epsilon^2} \right)^{1/2} \) [23].

Using Lemma 1 to estimate the right hand side of the above expression, with \( \beta = 2 \), \( T_{\alpha, \beta} = \frac{1}{\sqrt{\max_k(D_k)}} \), \( ||a_n||_\beta = \sqrt{N_R} \), \( \alpha = 2\sqrt{N_R \max_k(D_k)} \), we obtain \( \underline{U}_1 \leq K_1 \epsilon^2 p_0 + (1 - K_1 \epsilon^2) \) \( q/p - p_0 \) \( \leq \) \( K_1 \epsilon^2 (1 - K_1 \epsilon^2) \) \( N_r(N_r - 1) + 2 \) \( \exp\left( -\frac{\gamma^2}{\max_k(D_k)} \right) \). Choosing \( \epsilon \) to minimize the above upper bound, we find that \( \underline{U}_1 = O\left( \frac{\log(N_R/N_p)}{N_R} \right) \).

REFERENCES