OPTIMAL SIGNAL DESIGN FOR ESTIMATION OF CORRELATED MIMO CHANNELS

Jayesh H. Kotecha and Akbar M. Sayeed
Dept. of ECE, University of Wisconsin at Madison, Madison, WI 53706.
Email: *jkotecha@ece.wisc.edu, akbar@engr.wisc.edu.*

ABSTRACT

Optimal estimation of multi-input multi-output correlated channels using pilot signals is considered in this paper, assuming knowledge of the second order channel statistics at the transmitter. Assuming a block fading channel model and minimum mean square error (MMSE) estimation at the receiver, we design the transmitted signal to optimize two criteria: MMSE and the conditional mutual information between the MIMO channel and the received signal. Our analysis is based on the recently proposed virtual channel representation for uniform linear arrays, which corresponds to beamforming in fixed virtual directions and exposes the structure and the true degrees of freedom in correlated channels. However, the analysis can be generalized to other known channel models. We show that optimal signaling is in block form corresponding to beams transmitted in successive time intervals along the transmit virtual angles, with powers determined by water filling arguments based on the optimization criteria. The block length depends on the channel correlation and decreases with SNR. Consequently, from a channel estimation viewpoint, a faster fading rate can be tolerated at low SNRs relative to higher SNRs.

1. INTRODUCTION

Multi-antennae communications systems are gaining prominence due to the higher capacity and reliability they can afford [1],[2]. Often, an implicit assumption in the analysis is the accurate knowledge of the channel at the receiver. However, in practice the channel has to be estimated, typically using pilot symbols. In a rich scattering environment, the assumption of i.i.d. channels is valid and multi-input multi-output (MIMO) channel estimation can be done straightforwardly using for example least squares or MMSE techniques [3]. However, this idealized assumption does not necessarily hold and hence a study of correlated channels is of interest. In this work, we investigate transmit signal design for optimal estimation of correlated MIMO Rayleigh flat fading channels, assuming that the receiver and transmitter[1] have knowledge of the second order statistics of the MIMO channel[2]. This feedback information is exploited by the transmitter to optimize channel estimation errors at the receiver, where MMSE channel estimates are obtained. We design the transmit signal to satisfy one of two criteria : minimization of the MMSE at the receiver or maximization of the conditional mutual information between the channel and the received signal.

In [4], the virtual channel representation is proposed assuming uniform linear arrays (ULA) at the transmitter and receiver. The virtual representation characterizes the channel in the spatial domain by beamforming in the direction of fixed virtual angles determined by the spatial resolution of the arrays, which is analogous to representing the channel in beamspace or wavenumber domain. A MIMO channel with P transmit and Q receive antennae has a maximum of PQ unknowns to be estimated. However, correlated MIMO channels possess fewer degrees of freedom and hence fewer than PQ parameters need to be estimated. The non-vanishing and approximately uncorrelated elements of the virtual channel matrix represent the degrees of freedom in the channel. We develop our signal design based on the virtual representation. The techniques developed here however can be applied to more general channel representations like the one in [5].

We show that the optimal transmit signal is a block signal consisting of beams transmitted in succession along the active fixed transmit virtual angles, corresponding to directions in which scatterers are present. Equivalently, the (scattering) environment is scanned along the transmit virtual angles one by one to determine the presence of scattering clusters, by measuring the signals along the receive virtual angles for each transmitted beam. The power transmitted along the beams is determined by *water filling* arguments resulting from the two criteria under a finite power constraint. Power is possibly assigned to a beam only if the second order statistics indicate the presence of significant scattering in that direction. However, the power assigned to the transmit beams depends on the signal to noise ratio (SNR) as well. Specifically, at low SNR the strongest beam is assigned all the power. As SNR increases, the power is assigned to increasing number of beams depending on the channel covariance matrix[3].

2. MIMO CHANNEL MODEL

Consider a narrowband frequency non-selective MIMO channel with P transmit and Q receive antennae. With k indicating discrete time, if $\mathbf{s}(k)$ is the transmit vector of dimension P, then the Q dimensional received signal $\mathbf{x}(k)$ can be written as

$$\mathbf{x}(k) = \mathbf{H}(k)\mathbf{s}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{H}(k)$ is the $Q \times P$ channel gain matrix. $\mathbf{n}(k)$ is the Q dimensional zero mean, complex white Gaussian noise vector, with covariance matrix $\sigma^2 \mathbf{I}_Q$. The channel gain between the $n-th$ receive and $m-th$ transmit antenna is denoted by $\mathbf{H}[m,n]$.

In [4], the *virtual* channel representation is proposed where the transmit and receive antennae are uniform linear arrays (ULA). If

This research is supported in part by NSF Grant Nos. CCR-9875805 and CCR-0113385 and ONR Grant No. N00014-01-1-0825.

[1]This is often called covariance feedback.

[2]The assumption is reasonable, since the second order statistics are much less dynamic than the channel itself. Thus, they can be estimated more reliably and need to be updated less frequently.

[3]*Notation*: For an integer Q, $\mathbf{I}_Q$ is a $Q \times Q$ identity matrix. If $\mathbf{X}$ is a $Q \times K$ matrix, then it's lower case letter $\mathbf{x} = \mathrm{vec}(\mathbf{X})$ denotes the $QK \times 1$ vector obtained by stacking columns of $\mathbf{X}$. $\otimes$ denotes the Kronecker product. $\mathbf{X}^*, \mathbf{X}^T, \mathbf{X}^H$ denote the complex conjugate, transpose and hermitian of $\mathbf{X}$. The inverse and pseudo-inverse of $\mathbf{X}$ are denoted by $\mathbf{X}^{-1}$ and $\mathbf{X}^\dagger$. $\mathrm{tr}(\mathbf{X})$ denotes the trace of the square matrix $\mathbf{X}$. $\mathrm{diag}([a_1, \ldots, a_Q])$ is a $Q \times Q$ diagonal matrix with diagonal elements $a_1, \ldots, a_Q$. $E(\cdot)$ denotes the expectation operator. $(x)^+ = \max(0, x)$.

d_T and d_R are the transmit and receive array spacings, then $\mathbf{H}$ can be related to the physical propagation environment via the array steering and response vectors
$\mathbf{a}_T(\theta_T) = \frac{1}{\sqrt{P}}[1, \exp(-j2\pi\theta_T), \ldots, \exp(-j2\pi(P-1)\theta_T)]^T$,
$\mathbf{a}_R(\theta_R) = \frac{1}{\sqrt{Q}}[1, \exp(-j2\pi\theta_R), \ldots, \exp(-j2\pi(Q-1)\theta_R)]^T$,
where the θ is the delay between the signals received at adjacent elements in the array due to a point source at angle ϕ (relative to a horizontal axis). If λ is the wavelength of propagation, then $\theta = \frac{d}{\lambda}\sin\phi$. We will interpret θ as a normalized angle. The linear virtual channel representation in [4] exploits the finite dimensionality of the spatial signal space arising from finite number of array elements and finite array apertures. Without loss of generality, assume P and Q to be odd and define $\tilde{Q} = (Q-1)/2$ and $\tilde{P} = (P-1)/2$. The virtual channel representation is given by

$$\mathbf{H} = \sum_{q=-\tilde{Q}}^{\tilde{Q}} \sum_{p=-\tilde{P}}^{\tilde{P}} \mathbf{H}_V[p,q]\mathbf{a}_R(\tilde{\theta}_{R,q})\mathbf{a}_T^H(\tilde{\theta}_{T,p}) = \tilde{\mathbf{A}}_R \mathbf{H}_V \tilde{\mathbf{A}}_T^H$$

where $\tilde{\mathbf{A}}_R = [\mathbf{a}_R(\tilde{\theta}_{R,-\tilde{Q}}), \ldots, \mathbf{a}_R(\tilde{\theta}_{R,+\tilde{Q}})]$ $(Q \times Q)$ and $\tilde{\mathbf{A}}_T = [\mathbf{a}_T(\tilde{\theta}_{T,-\tilde{Q}}), \ldots, \mathbf{a}_T(\tilde{\theta}_{T,+\tilde{Q}})]$ $(P \times P)$ are defined by the *fixed* virtual angles $\tilde{\theta}_{R,q}$ and $\tilde{\theta}_{T,p}$ and are full rank. We assume that the spatial virtual angles are uniformly spaced [4] and hence $\tilde{\mathbf{A}}_T$ and $\tilde{\mathbf{A}}_R$ are discrete Fourier transform matrices (and hence unitary). Note that the virtual model is *linear* in the gains and spatial angles, since these angles are fixed a priori. Note that we can write $\mathbf{h} = \mathrm{vec}(\mathbf{H}) = (\tilde{\mathbf{A}}_T^* \otimes \tilde{\mathbf{A}}_R)\mathbf{h}_V$. The resulting channel correlation has a *Kronecker* structure given by

$$\mathbf{R} = E(\mathbf{h}\mathbf{h}^H) = (\tilde{\mathbf{A}}_T^* \otimes \tilde{\mathbf{A}}_R)\mathbf{R}_V(\tilde{\mathbf{A}}_T^* \otimes \tilde{\mathbf{A}}_R)^H. \quad (2)$$

An important consequence of the virtual modelling is that, the elements of $\mathbf{H}_V$ are approximately *uncorrelated* and hence $\mathbf{R}_V$ is approximately diagonal regardless of the correlation structure of $\mathbf{R}$ [4]. The structure obtained by the virtual model allows simplification in signal design and provides interesting interpretations as shall be seen.

The techniques developed in this paper can be straightforwardly generalized to channels where the channel matrix can be expressed as

$$\mathbf{H} = \mathbf{U}_R \mathbf{H}_V \mathbf{U}_T^H \quad (3)$$

where $\mathbf{U}_T$ and $\mathbf{U}_R$ are the transmit and receive unitary matrices and the elements of $\mathbf{H}_V$ are *uncorrelated* but not necessarily identically distributed. The resulting channel correlation has a Kronecker structure similar to (2). Such channel models may arise as a consequence of the array geometry as was seen above in the case of ULAs. An example is the channel model, where it is assumed that the transmitter and receiver antennae arrays have correlated elements [5]. The channel matrix can be written as

$$\mathbf{H} = \mathbf{\Sigma}_R^{1/2}\mathbf{H}_w\mathbf{\Sigma}_T^{1/2} = \mathbf{U}_R \mathbf{H}_V \mathbf{U}_T^H \quad (4)$$

where the elements of $\mathbf{H}_w$ are i.i.d. The matrices $\mathbf{\Sigma}_T$ and $\mathbf{\Sigma}_T$ are the transmit and receive array correlation matrices with eigen value decompositions (EVD) $\mathbf{U}_T\Lambda_T\mathbf{U}_T^H$ and $\mathbf{U}_R\Lambda_R\mathbf{U}_R^H$ respectively. The elements of $\mathbf{H}_V$ are uncorrelated with diagonal covariance matrix given by $\mathbf{R}_V = \Lambda_T \otimes \Lambda_R$ [6].

Since $\mathbf{H}$ and $\mathbf{H}_V$ are unitarily equivalent, estimation of the MIMO channel can be equivalently performed by obtaining estimates of $\mathbf{H}_V$. From (1) and (3), we can write the received signal as

$$\mathbf{x}(k) = \tilde{\mathbf{A}}_R \mathbf{H}_V(k) \tilde{\mathbf{A}}_T^H \mathbf{s}(k) + \mathbf{n}(k). \quad (5)$$

In the eigen or virtual domain,

$$\mathbf{x}_V(k) = \mathbf{H}_V(k)\mathbf{s}_V(k) + \mathbf{n}_V(k) \quad (6)$$

where $\mathbf{x}_V = \tilde{\mathbf{A}}_R^H\mathbf{x}$ and $\mathbf{s}_V = \tilde{\mathbf{A}}_T^H\mathbf{s}$ are the projections of the received and transmitted signals onto the fixed receive and transmit response vectors respectively. Equation (6) provides an interesting interpretation of transmission in the virtual domain. Each element of $\mathbf{x}_V$ ($\mathbf{s}_V$) corresponds to a signal received (transmitted) from (to) the fixed virtual angles $\tilde{\theta}_{T,p}$ ($\tilde{\theta}_{R,q}$) and the corresponding element in $\mathbf{H}_V$ indicates the coupling gain between these angles [4]. Note that since $\tilde{\mathbf{A}}_R$ is unitary, $\mathbf{n}_V = \tilde{\mathbf{A}}_R\mathbf{n}$ is zero mean, white Gaussian with covariance $\sigma^2\mathbf{I}_q$.

In the following development, we assume the MIMO channel to be block fading, i.e. $\mathbf{H}(k) = \mathbf{H}$ for $k = 1, \ldots, K$ and the channel is independent between different blocks of K symbols. Assuming that training symbols $\mathbf{s}(k)$, $k = 1, \ldots, K$ are sent in a block mode and denoting $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(K)]$, the block fading model is given by

$$\mathbf{X}_V = \mathbf{H}_V\mathbf{S}_V + \mathbf{N}_V,$$

where $\mathbf{X}_V = [\mathbf{x}_V(1), \ldots, \mathbf{x}_V(K)]$, $\mathbf{S}_V = [\mathbf{s}_V(1), \ldots, \mathbf{s}_V(K)]$ and $\mathbf{N}_V = [\mathbf{n}_V(1), \ldots, \mathbf{n}_V(K)]$. Stacking the columns of $\mathbf{X}_V$, we obtain

$$\tilde{\mathbf{x}}_V = \mathrm{vec}(\mathbf{X}_V) = (\mathbf{S}_V^T \otimes \mathbf{I}_Q)\mathrm{vec}(\mathbf{H}_V) + \mathrm{vec}(\mathbf{N}_V) = \tilde{\mathbf{S}}_V\mathbf{h}_V + \tilde{\mathbf{n}}_V \quad (7)$$

where we denote $\tilde{\mathbf{S}}_V = (\mathbf{S}_V^T \otimes \mathbf{I}_Q)$. Using (7), we proceed with the estimation of $\mathbf{h}_V$, which is a PQ vector. Clearly, since the maximum number of unknowns in $\mathbf{h}_V$[4] is PQ, we need to transmit a block of $K \le P$ symbols [7]. Hence, we need the quasi-static channel to be constant for only $K \le P$ time periods.

3. MMSE AND MAP ESTIMATION

The model (7) is linear in $\mathbf{h}_V$ and Gaussian. Hence, it can be shown that the linear MMSE estimate, the MMSE estimate and the MAP estimate are *identical*. In this paper, we assume that the covariance matrix $\mathbf{R}_V = E(\mathbf{h}_V\mathbf{h}_V^H)$ (or equivalently $\mathbf{R} = E(\mathbf{h}\mathbf{h}^H)$) is known. The linear MMSE estimator minimizes the error $\mathrm{MSE} = E[\|\mathbf{h}_V - \hat{\mathbf{h}}_V\|^2]$. The resulting linear estimate is

$$\hat{\mathbf{h}}_V = \mathbf{G}_{opt}\tilde{\mathbf{x}}_V \quad (8)$$

where $\mathbf{G}_{opt}$ is a $PQ \times PQ$ matrix given by

$$\mathbf{G}_{opt} = \arg\min_{\mathbf{G}} E[\|\mathbf{h}_V - \mathbf{G}\tilde{\mathbf{x}}_V\|^2] = \mathbf{R}_V\tilde{\mathbf{S}}_V^H(\tilde{\mathbf{S}}_V\mathbf{R}_V\tilde{\mathbf{S}}_V^H + \sigma^2\mathbf{I})^{-1}. \quad (9)$$

Using the orthogonality principle, the error covariance matrix and the minimum MSE are

$$\mathbf{C}_e = \mathbf{R}_V - \mathbf{R}_V\tilde{\mathbf{S}}_V^H(\tilde{\mathbf{S}}_V\mathbf{R}_V\tilde{\mathbf{S}}_V^H + \sigma^2\mathbf{I})^{-1}\tilde{\mathbf{S}}_V\mathbf{R}_V$$
$$= (\mathbf{R}_V^\dagger + \frac{1}{\sigma^2}\tilde{\mathbf{S}}_V^H\tilde{\mathbf{S}}_V)^{-1}, \quad (10)$$

$$\mathrm{MMSE} = \mathrm{tr}(\mathbf{C}_e(\tilde{\mathbf{S}}_V)). \quad (11)$$

respectively. The conditional mutual information (CMI) between the received signal and the channel $\mathbf{h}_V$ is given by

$$\mathrm{CMI}(\tilde{\mathbf{S}}_V) = \log\det(\mathbf{I} + \frac{1}{\sigma^2}\tilde{\mathbf{S}}_V\mathbf{R}_V\tilde{\mathbf{S}}_V^H). \quad (12)$$

[4]The number of unknowns in $\mathbf{h}_V$ would be smaller in correlated channels. If the prior variance of a given element of $\mathbf{h}_V$ is zero, then it implies that the element is itself zero.

4. OPTIMUM SIGNAL DESIGN

We consider the design of the optimum transmit block signal $\tilde{\mathbf{S}}_V$ (or equivalently $\mathbf{S}_V$) with respect to two criteria: minimization of the MMSE (11) and maximization of the mutual information (12) between the channel and received signal conditioned on the transmitted block signal. We state the two optimization problems as follows:

$$\min_{\tilde{\mathbf{S}}_V} \ \mathrm{tr}(\mathbf{R}_V^{\dagger} + \frac{1}{\sigma^2}\tilde{\mathbf{S}}_V^H \tilde{\mathbf{S}}_V)^{-1} \quad \text{s.t.} \quad \mathrm{tr}(\tilde{\mathbf{S}}_V^H \tilde{\mathbf{S}}_V) \leq P\beta, \quad (13)$$

$$\max_{\tilde{\mathbf{S}}_V} \ \log\det(\mathbf{I} + \frac{1}{\sigma^2}\tilde{\mathbf{S}}_V \mathbf{R}_V \tilde{\mathbf{S}}_V^H) \ \text{s.t.} \ \mathrm{tr}(\tilde{\mathbf{S}}_V^H \tilde{\mathbf{S}}_V) \leq P\beta, \quad (14)$$

where β is the total transmitted power and P is the number of transmit antennae. Note that the constraint $\mathrm{tr}(\tilde{\mathbf{S}}_V^H \tilde{\mathbf{S}}_V) \leq P\beta$ is equivalent to the finite power constraint $\mathrm{tr}(\mathbf{S}_V^H \mathbf{S}_V) = \mathrm{tr}(\mathbf{S}^H \mathbf{S}) \leq \beta$. We develop the signal design using the SVD of the transmitted block matrix. Denote the SVDs of $\mathbf{S}_V^T = \mathbf{U}_S \Lambda_S \mathbf{V}_S^H$ and $\tilde{\mathbf{S}}_V = \mathbf{U}_{\tilde{S}} \Lambda_{\tilde{S}} \mathbf{V}_{\tilde{S}}^H$, where $\mathbf{U}_S, \mathbf{V}_S, \mathbf{U}_{\tilde{S}}$ and $\mathbf{V}_{\tilde{S}}$ are unitary matrices and Λ_S and $\Lambda_{\tilde{S}}$ are diagonal matrices. Since $\tilde{\mathbf{S}}_V = (\mathbf{S}_V^T \otimes \mathbf{I}_Q)$, it follows that $\mathbf{U}_{\tilde{S}} = \mathbf{U}_S \otimes \mathbf{I}_Q$, $\Lambda_{\tilde{S}} = \Lambda_S \otimes \mathbf{I}_Q$ and $\mathbf{V}_{\tilde{S}} = \mathbf{V}_S \otimes \mathbf{I}_Q$. The following theorem states our main result, for a proof see [6].

Theorem 1 *Consider the constrained optimization problems in (13) and (14) respectively. The globally optimal solution has a structure given by*

$$\tilde{\mathbf{S}}_{V,opt} = \tilde{\Lambda}_{opt} \mathbf{V}_{\tilde{S}} \quad (15)$$

where $\tilde{\Lambda}_{opt} \in \mathbb{R}^{PQ \times PQ}$. *The optimal* $\mathbf{V}_{\tilde{S}}$ *is a matrix of the eigenvectors of* $\mathbf{R}_V$, *i.e* $\mathbf{V}_{\tilde{S}} = \mathbf{I}$ *and* $\tilde{\Lambda}_{opt}$ *is the solution to*

$$\tilde{\Lambda}_{opt} = \tilde{\Lambda}_M = \arg\min_{\Lambda_{\tilde{S}}} \mathrm{tr}(\mathbf{R}_V^{\dagger} + \frac{1}{\sigma^2}\Lambda_{\tilde{S}}^H \Lambda_{\tilde{S}})^{-1} \quad (16)$$
$$\text{s.t.} \quad \mathrm{tr}(\Lambda_{\tilde{S}}^H \Lambda_{\tilde{S}}) \leq P\beta$$

$$and \ \ \tilde{\Lambda}_{opt} = \tilde{\Lambda}_C = \arg\max_{\Lambda_{\tilde{S}}} \log\det(\mathbf{I} + \Lambda_{\tilde{S}} \mathbf{R}_V \Lambda_{\tilde{S}}^H) \quad (17)$$
$$\text{s.t.} \quad \mathrm{tr}(\Lambda_{\tilde{S}}^H \Lambda_{\tilde{S}}) \leq P\beta$$

respectively.

Since $\tilde{\mathbf{S}}_V = (\mathbf{S}_V^T \otimes \mathbf{I}_Q)$, from Theorem 1 the optimal transmit signal is $\mathbf{S}_V = \Lambda_{opt}$ where Λ_{opt} is the solution to

$$\arg\min_{\Lambda_S} \sum_{i=1}^{P}\sum_{j=1}^{Q} \left(\frac{\sigma^2 \mathbf{R}_V[(i-1)Q+j,(i-1)Q+j]}{\sigma^2 + \mathbf{R}_V[(i-1)Q+j,(i-1)Q+j]\beta_i} \right) \quad (18)$$

$$\arg\max_{\Lambda_S} \sum_{i=1}^{P}\sum_{j=1}^{Q} \log\left(1 + \frac{\mathbf{R}_V[(i-1)Q+j, \ (i-1)Q+j]\beta_i}{\sigma^2}\right) \quad (19)$$

subject to the constraint $\sum_{i=1}^{P}\beta_i \leq \beta$, $\beta_i = |\Lambda_{opt}(i,i)|^2$, for the MMSE and CMI criteria respectively. Thus, the optimal transmit signal is a block diagonal signal (in the virtual domain). The optimal signal structure specifies that during the P block transmission, at each time instant $i \in 1,\ldots,P$, the signal is transmitted along the i-th transmit eigen vector with the powers specified by β_i. Due to the diagonal structure of $\mathbf{S}_V$, $\mathbf{G}_{opt}$ (9) and $\mathbf{C}_e$ (10) become diagonal, which enables *independent* processing at the receiver. The channel estimate is given by

$$\hat{\mathbf{h}}_V((i-1)Q+j) = \left(\frac{\mathbf{R}_V[(i-1)Q+j,(i-1)Q+j] \ \Lambda_{opt}^H(i,i)}{\sigma^2 + \mathbf{R}_V[(i-1)Q+j,(i-1)Q+j] \ |\Lambda_{opt}(i,i)|^2}\right) \cdot \mathbf{x}_V((i-1)Q+j),$$

for $j = 1,\ldots,Q; \ i = 1,\ldots,P$. From this equation, note that the i-th transmission allows us to estimate the Q elements in the $i-th$ column of $\mathbf{H}_V$, i.e ($\mathbf{h}_V((i-1)Q+1),\ldots,\mathbf{h}_V((i-1)Q+Q)$). During the block transmission, the scattering environment is scanned sequentially to estimate each column of $\mathbf{H}_V$.

4.1. Water-filling solution

The constrained nonlinear optimizations in (18) and (19) are the so called 'water-filling' problems and can be solved using Langrange multipliers and using the Kuhn-Tucker conditions to verify that the solutions are non-negative. However, for the general case of P transmit and Q receive antennae, we have not been able to find a closed form solution and hence it has to be obtained numerically. In the following, we obtain approximate closed form solutions in the low SNR and high SNR regions to obtain some insight. Closed form solutions exist for the special cases of a MISO channel [8] and the transmit and receive correlated channel (4) where either Σ_T or Σ_R is equal to $\sigma^2\mathbf{I}$, for details see [6].

For the following discussion, we define the transmitted signal to noise ratio (TSNR) as the ratio of the transmitted signal power to the noise power $\frac{\beta}{\sigma^2}$ and the received signal to noise ratio (RSNR) between the $i-th$ transmit and $j-th$ receive angle pair as the ratio of the received signal power to the noise power $\mathrm{RSNR}(i,j) = \frac{\mathbf{R}_V[(i-1)Q+j,(i-1)Q+j]\beta_i}{\sigma^2}$ for $i = 1,\ldots,P; \ j = 1,\ldots,Q$.

Consider the high RSNR case, where $\mathrm{RSNR}(i,j) \gg 1$. In the following discussion, denote elements for which the high RSNR condition is true as 'active' and columns which have at least one active element as active columns. Let Q_i be the number of active elements in the i-th column (or equivalently the number of active receive elements the i-th transmit beam couples with). Using Langrange multipliers, it can be shown that for high RSNR case, MMSE and CMI criteria assign power according to

$$\text{MMSE}: \quad \beta_i = \frac{\sqrt{Q_i}\beta}{\sum_{i=1}^{P}\sqrt{Q_i}} \quad i = 1,\ldots,P,$$

$$\text{and} \ \ \text{CMI}: \quad \beta_i = \frac{Q_i\beta}{\sum_{i=1}^{P}Q_i} \quad i = 1,\ldots,P,$$

respectively. Thus, the CMI (MMSE) criterion assigns power to the transmit beams in proportion to the sum (square root of the sum) of the active elements they couple with at the receiver. In the extreme case, when all the elements of $\mathbf{H}_V$ are active, then *equal* power is distributed at all transmit branches for both the criteria.

Consider the low RSNR case, where $\mathrm{RSNR}(i,j) \ll 1, \ \forall i,j$. Using Langrange multipliers, it can be shown that the MMSE and CMI criteria assign all the power β to the $k-$th transmit beam such that

$$\text{MMSE}: \quad k = \arg\max_i \sum_{j=1}^{Q} \mathbf{R}_V^2[(i-1)Q+j,(i-1)Q+j],$$

$$\text{and} \ \ \text{CMI}: \quad k = \arg\max_i \sum_{j=1}^{Q} \mathbf{R}_V[(i-1)Q+j,(i-1)Q+j],$$

respectively. Thus at low RSNR, the CMI (MMSE) criterion assigns all the power to that transmit angle for which the sum (sum of squares) of the variances of the corresponding virtual receive elements is maximum. From the extreme cases, we conclude that the number of transmit beams to be sent and hence the block length K depends on the SNR. For medium SNR, $1 \leq K \leq P$ and the powers will be determined by the water filling criteria. Also note that for i.i.d. channels, equal power will be assigned to all transmit beams irrespective of the SNR.

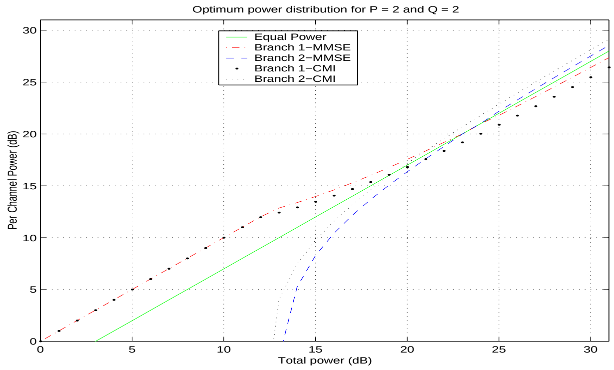

Fig. 1. Optimal power distribution for $P = Q = 2$ & $\mathbf{R}_V = \mathbf{R}_V^{(1)}$.

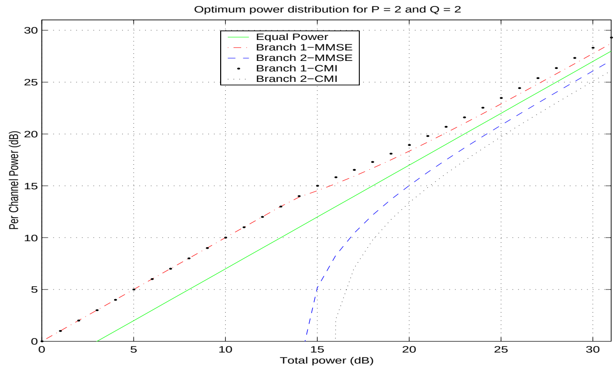

Fig. 2. Optimal power distribution for $P = Q = 2$ & $\mathbf{R}_V = \mathbf{R}_V^{(2)}$.

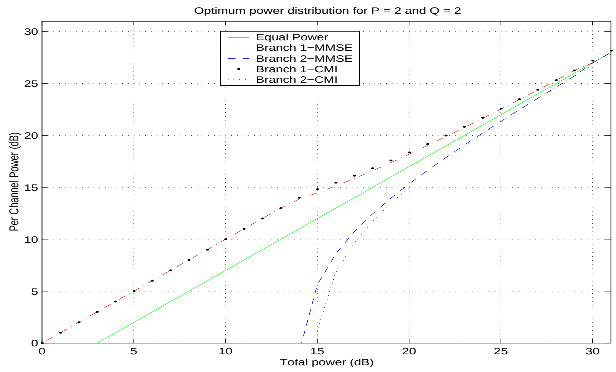

Fig. 3. Optimal power distribution for $P = Q = 2$ & $\mathbf{R}_V = \mathbf{R}_V^{(3)}$.

5. INTERPRETATION AND SIMULATIONS

The optimal signal is a block of length $K \leq P$ and has a diagonal structure given by $\mathbf{S}_V = \Lambda_S$. The block $\mathbf{S}_V$ represents beams transmitted in succession along the fixed virtual transmit angles, with the powers given by the water filling arguments (18) and (19) for the MMSE and CMI criteria respectively. Basically, the scattering environment is scanned along the virtual transmit angles one by one, to determine the presence of scatterers, by measuring the signal along the receive virtual angles for each transmitted beam. The i-th transmitted beam is used to determine the i-th column of $\mathbf{H}_V$. Depending on $\mathbf{R}_V$ and the SNR, power is assigned to the beams by water filling, which identifies the *active set* of virtual transmit angles. Hence the block length K, which is exactly equal to the size of this active set, depends on the SNR and $\mathbf{R}_V$. In particular, for low SNR, $K = 1$, while for high SNR K has a maximum value equal to the number of active columns determined from $\mathbf{R}_V$ (which is a maximum of P) and for medium SNR, $1 \leq K \leq P$. This in turn implies that at low SNR, a faster *fading rate* can be tolerated than at high SNR, since fewer essential parameters need to be estimated. For high SNR, the CMI (MMSE) criterion assigns the power to the transmit angles in proportion to the sum ($\sqrt{\text{sum}}$) of the active elements they couple with at the receiver. As the SNR decreases, the weakest transmit beam (as determined by the water filling criteria) is dropped. As the SNR decreases, this process continues until finally the CMI (MMSE) criterion assigns all the power to the strongest transmit beam, i.e one for which the sum (sum of squares) of the variances of the corresponding virtual receive elements is maximum. This is illustrated in Figures 1 - 3. In all figures, the total TSNR (in dB) along the x-axis is given by $10\log_{10}(\beta/\sigma^2)$, while the y-axis shows the branch TSNR in dB given by $10\log_{10}(\beta_i/\sigma^2)$. Powers are show for the two transmit angles for the MMSE and CMI criteria and the equal power assignment is also plotted for comparison. Figure 1 shows the power assignments for the MIMO case with $P = Q = 2$ and covariance matrix is given by $\mathbf{R}_V^{(1)} = \text{diag}([1 \;\; 0 \;\; 0.01 \;\; 0.05])$, where the first two elements are the variances of the elements in the first column of $\mathbf{H}_V$ and the next two are those of the second column. Observe that for high SNR, the second transmit beam gets 66% (58%) power according to CMI (MMSE) criterion. The powers are reversed in Figure 2 where $\mathbf{R}_V^{(2)} = \text{diag}([1 \;\; 0.1 \;\; 0 \;\; 0.05])$. Finally in Figure 3 where $\mathbf{R}_V^{(3)} = \text{diag}([1 \;\; 0.1 \;\; 0.01 \;\; 0.05])$, at high SNR both branches get equal power. In all cases, as SNR decreases the weaker beam is dropped and the stronger beam gets all the power.

6. REFERENCES

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *AT & T Bell Labs Internal Technical Memo*, June 1995.

[2] G. J. Foschini, "Layered space-time architechture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[3] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Proc. Magazine*, pp. 49–83, Nov. 1997.

[4] A. M. Sayeed, "Deconstructing multi-antennae fading channels," *IEEE Transactions of Signal Processing*, Oct. 2002.

[5] D. Shiu, G. Foschini, M. Gans, and J. Kahn, "Fading correlation and its effect on the capacity of multielement antenna systems," *IEEE Transactions on Communications*, Mar. 2000.

[6] J. H. Kotecha and A. M. Sayeed, "Optimal Estimation of Correlated MIMO channels," *http://dune.ece.wisc.edu/∼akbar/pub.html*, preprint, 2002.

[7] B. Hassibi and B. Hochwald, "Optimal training in space-time systems," *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers*, vol. 1, pp. 743 –747, 2000.

[8] G. G. Giannakis and S. Zhou, "Optimal transmit-diversity precoders for random fading channels," *Proc. of Globecomm*, pp. 1839–1843, 2000.