Tracking a MIMO Channel Singular Value Decomposition via Projection Approximation

Sharon Gannot[†], Amir Leshem[†], Ofer Shayevitz[§] and Ephi Zehavi[†]

***Abstract*— A bidirectional multiple-input multiple-output (MIMO) time varying channel is considered. The projection approximation subspace tracking (PAST) algorithm is used on both terminals in order to track the singular value decomposition of the channel matrix. Simulations using an autoregressive channel model and also a sampled MIMO indoor channel are performed, and the expected capacity degradation due to the estimation error is evaluated.**

I. Channel Model

The MIMO channel model with N_t transmit antennas and N_r receive antennas is given by

$$\boldsymbol{y}_n = H_n \boldsymbol{x}_n + \boldsymbol{z}_n \tag{1}$$

Where $\boldsymbol{x}_n \in \mathbb{C}^{N_t \times 1}$ is the channel's input vector, $\boldsymbol{y}_n \in \mathbb{C}^{N_r \times 1}$ is the channel's output vector, $H_n \in \mathbb{C}^{N_r \times N_t}$ is the channel matrix, and $\boldsymbol{z}_n$ is a circular symmetric complex white Gaussian noise, all given at time n.

We consider a bidirectional link, and assume that the gains of the RF chains on both terminals are fully compensated for, so that the channel matrices in both directions reflect only the effect of the physical link itself. Under this *reciprocity* assumption, the reverse channel model is similarly given by

$$\tilde{\boldsymbol{y}}_n = H_n^\dagger \tilde{\boldsymbol{x}}_n + \tilde{\boldsymbol{z}}_n$$

where $H_n^\dagger$ is the conjugate transpose of H_n.

In our simulations, we assume the following fading model for the channel matrix H_n:

$$H_n = \sum_{k=1}^{P} \alpha_k H_{n-k} + W_n$$

where $\{\alpha_k\}_{k=1}^P$ are the coefficients of the above autoregressive (AR) model of order P normalized to a unity DC gain, and the elements of the innovation matrix W_n are i.i.d $\sim \mathcal{CN}(0,1)$.

II. Tracking Algorithm

The channel matrix H_n can be factored using the singular value decomposition (SVD) as follows:

$$H_n = U_n S_n V_n^\dagger \tag{2}$$

where $U_n \in \mathbb{C}^{N_r \times N_r}$ and $V_n \in \mathbb{C}^{N_t \times N_t}$ are unitary matrices of the the left and right singular vectors of H_n respectively, and $S_n \in \mathbb{C}^{N_r \times N_t}$ is a diagonal matrix having the nonnegative $\{s_k\}_{k=1}^{\min(N_r,N_t)}$ singular values of H_n on its diagonal. Consequently, the SVD factorization of $H_n^\dagger$ is given by

$$H_n^\dagger = V_n S_n^T U_n^\dagger \tag{3}$$

If the matrix V_n is known on the forward transmission site at time n, and U_n on the reverse site, then the channel in both directions can be diagonalized:

$$\begin{aligned}(U_n^\dagger \boldsymbol{y_n}) &= S_n(V_n^\dagger \boldsymbol{x_n}) + U_n^\dagger \boldsymbol{z_n}\\ (V_n^\dagger \tilde{\boldsymbol{y_n}}) &= S_n^T(U_n^\dagger \tilde{\boldsymbol{x_n}}) + V_n^\dagger \tilde{\boldsymbol{z_n}}\end{aligned}$$

and if the singular values are also known (i.e., the channel matrix is essentially known), communications in both directions narrows down to parallel independent additive white Gaussian noise (AWGN) channels. If the channel matrix remains constant then capacity is attained via Gaussian water filling over the singular vectors, and otherwise power control strategies must be used. We shall restrict our capacity discussion however to a Gaussian codebook with equal constant power per singular mode, and also assume that the singular values are compensated for by the decoder (e.g., in a decision directed manner).

[†] The authors are with the School of Electrical Engineering, Bar-Ilan University, Israel

[§] The author is with the department of Electrical Engineering - Systems, Tel-Aviv University, Israel

Following this, the information transmitted in both directions is Gaussian distributed with zero mean, satisfying

$$E(\boldsymbol{x}_n \boldsymbol{x}_m^\dagger) = \rho I_{N_t}\, \delta_{m,n}\,, \quad E(\tilde{\boldsymbol{x}}_n \tilde{\boldsymbol{x}}_m^\dagger) = \tilde{\rho} I_{N_r}\, \delta_{m,n} \tag{4}$$

where I_M is the identity $M \times M$ matrix and $\delta_{m,n}$ is Kronecker's delta.

As we have seen, the forward site needs to know only the matrix V_n and the reverse site only the matrix U_n. Assume that an accurate estimate of V_n, U_n is given to the respective terminals. If the channel remains constant, then reliable communications may be attained via SVD. On a varying channel, one needs to keep track the variations in the matrices above. We shall therefore use the received data on each terminal in order to update the estimate of the corresponding matrix. We will only describe the estimation process for U_n performed in the reverse site using the data received from the forward link, as the estimate of V_n is similar. From now on we shall also assume for clarity that $N_t \leq N_r$.

The correlation matrix of the received vector is given by

$$\begin{aligned} R_{\boldsymbol{y}}(n) &\stackrel{\triangle}{=} E(\boldsymbol{y_n} \boldsymbol{y_n}^\dagger) = H_n E(\boldsymbol{x_n} \boldsymbol{x_n}^\dagger) H_n^\dagger + I_{N_r} \\ &= \rho H_n H_n^\dagger + I_{N_r} = U_n(\rho S_n S_n^T + I_{N_r}) U_n^\dagger \end{aligned}$$

thus the eigenvectors of $R_{\boldsymbol{y}}(n)$ are the left singular vectors of H_n, i.e., the columns of U_n. Therefore, we wish to estimate the eigenvectors of $R_{\boldsymbol{y}}(n)$ from the received data vectors $\boldsymbol{y}_n$.

The projection approximation subspace tracking (PAST) algorithm provides a convenient recursive method for estimating the subspace spanned by $R_{\boldsymbol{y}}(n)$. The method is based on the interpretation of that subspace as a solution to a minimization problem, and on the application of recursive least squares techniques combined with a subspace projection approximation. However, this is not satisfactory for our purposes since we are interested in the specific basis of orthonormal eigenvectors of $R_{\boldsymbol{y}}(n)$. Therefore, we will use a variant of PAST which is called PASTd, as it uses a deflation technique in order to provide the required orthonormal basis. Notice that the singular vectors in either the forward site or the backward site (or both, depending on the number of antennas in each site) span the entire vector space. Nevertheless, it is guaranteed that the obtained singular vectors constitute an orthonormal basis. In our context, the algorithm used on the receiver site is summarized below (vectors are in bold):

Set $d_0^i = 1\, \forall i$, Initialize $\boldsymbol{u}_0^i = (\overbrace{0,\ldots,0}^{i-1}, 1, 0, \ldots, 0)$

For n=1,2,...

$\boldsymbol{y}_n^i = \boldsymbol{y}_n$

For i=1,2,... N_r

$w_n^i = (\boldsymbol{u}_{n-1}^i)^\dagger \boldsymbol{y}_n$

$d_n^i = \beta d_n^i + |w_n^i|^2$

$\boldsymbol{e}_n^i = \boldsymbol{y}_n^i - \boldsymbol{u}_{n-1}^i w_n^i$

$\boldsymbol{u}_n^i = \boldsymbol{u}_{n-1}^i + \boldsymbol{e}_n^i \dfrac{(w_n^i)^*}{d_n^i}$

$\boldsymbol{y}_n^{i+1} = \boldsymbol{y}_n^i - \boldsymbol{u}_n^i w_n^i$

$0 < \beta \leq 1$ is the "forgetting factor" of the recursive scheme, and $\boldsymbol{u}_0^i$ are set to be the initial estimates to the left singular column vectors. The estimate $\widehat{U}_n$ of the matrix U_n at time n is given by

$$\widehat{U}_n = \left[\boldsymbol{u}_n^1\ \boldsymbol{u}_n^2\ \cdots \boldsymbol{u}_n^{N_r}\right] \tag{5}$$

and is guaranteed to be "almost orthonormal".

Similarly, the estimate $\widehat{V}_n$ of V_n is evaluated on the forward transmission site, using the data received from the reverse channel. The estimates $\widehat{V}_n, \widehat{U}_n$ are used for both transmission and reception on both sides. Hence, the effective channel matrix viewed by the forward channel receiving side is given by

$$\widehat{S}_n = \widehat{U}_n^\dagger U_n S_n V_n^\dagger \widehat{V}_n \tag{6}$$

which results in a diagonal channel had the estimate been exact.

The maximal (forward) rate achievable by our transmission scheme with full channel state information (CSI), i.e., with H_n known, is given by

$$C_n^{max} = \log\det\left(I_{N_t} + \rho S_n S_n^\dagger\right) \tag{7}$$

In practice, since the resulting channel is not diagonal, there are crossover effects between the modes, and in addition, there may be some noise enhancement since the estimates $\widehat{U}_n, \widehat{V}_n$ are not precisely orthonormal, although this effect is rather negligible.

For any matrix A we denote by $\text{diag}(A)$ a square diagonal matrix with the same diagonal as A. The effective maximal rate achievable in our setting is

$$\Lambda_n \triangleq \mathrm{diag}(\widehat{U}_n^\dagger \widehat{U}_n)$$
$$\Lambda_c \triangleq \mathrm{diag}(\widehat{S}_n \widehat{S}_n^\dagger) - \mathrm{diag}(\widehat{S}_n)\mathrm{diag}(\widehat{S}_n)^\dagger$$
$$\Psi \triangleq \mathrm{diag}(\widehat{S}_n)\mathrm{diag}(\widehat{S}_n)^\dagger$$
$$C_n^{eff} = \log\det\left(I_{N_t} + \rho\Psi\left(\Lambda_n + \rho\Lambda_c\right)^{-1}\right) \quad (8)$$

$\widehat{U}_n^\dagger \widehat{U}_n$ is the noise covariance matrix, and we are only concerned with its diagonal since the receiver assumes a parallel channel and disregards cross-correlations. The second expression inside the inverting parenthesis represents the interference from other modes, and the expression outside the parenthesis is the mode's power.

III. Simulations

Simulations were performed using an auto-regressive (AR) fading model for the channel matrix. The average capacity loss of our scheme relative to a scheme with full CSI was calculated for different AR parameters and different values of β, the "forgetting factor" of the PASTd algorithm. An average loss of less than 10% was attained for a reasonable fading process, with optimal β. Simulations were also performed for a sampled real indoor model[1], resulting in a loss not exceeding 3% for an optimal selection of β. Furthermore, capacity-versus-outage curves for both the AR and real indoor fading models are provided.

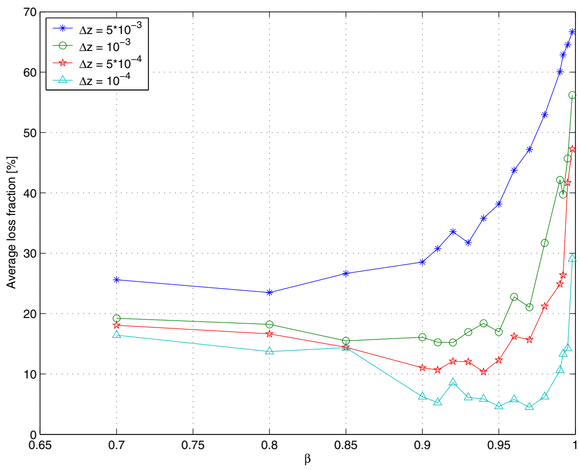

Fig. 1. β versus $\mathbb{E}(\frac{C^{max}-C^{eff}}{C^{max}})$ for AR(1) channel model with response $\frac{K}{1+(1-\Delta z)z^{-1}}$

[1]We are grateful to Metalink for providing the measurement

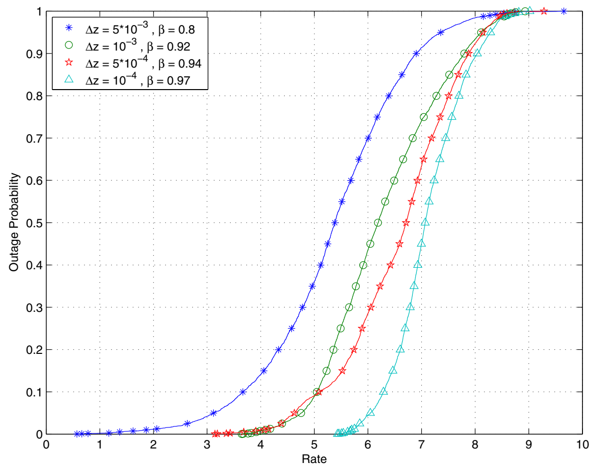

Fig. 2. Outage probability vs. rate with optimal β for AR(1) channel

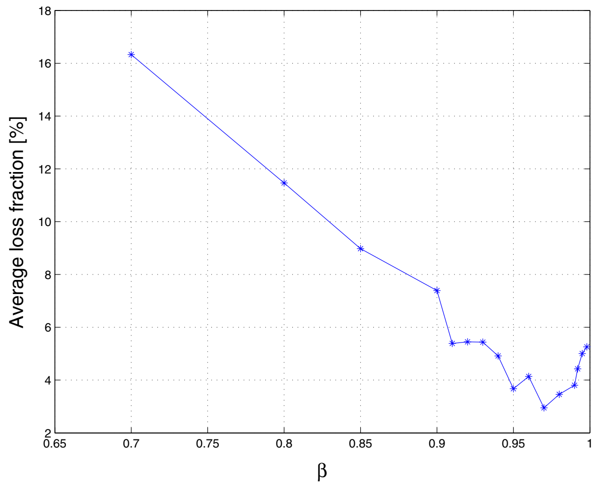

Fig. 3. β versus $\mathbb{E}(\frac{C^{max}-C^{eff}}{C^{max}})$ for sampled indoor channel

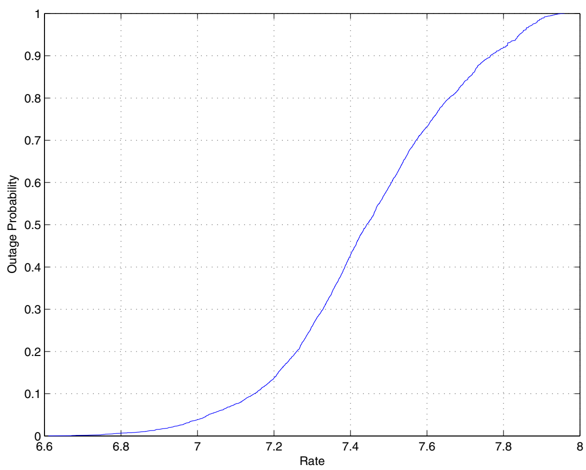

Fig. 4. Outage probability vs. rate with optimal β for indoor channel

REFERENCES

[1] B. Yang, "Projection approximation subspace tracking," *IEEE Trans. on signal proccessing,* Vol. 43, No. 1, pp. 95-107, Jan. 1995.

[2] G. Lebrun, J. Gao and M. Faulkner "MIMO Transmission over a time-varying channel using SVD," *IEEE Trans. on wireless communications,* Vol. 4, No. 2, pp. 757-764, Mar. 2005.

[3] I.E. Telatar, "Capacity of multi-antenna Gaussian channels", Tech. Rep. AT&T Bell Labs, 1995.

[4] T. Dahl, N. Christopherson and D. Gesbert, "Blind MIMO Eigenmode Transmission Based on the Algebraic Power Method," *IEEE Trans. on Signal Processing,* Vol. 52, N0. 9, pp. 2424-2431, Sep. 2004.