MIMO TRANSMISSION SUBSPACE TRACKING WITH LOW RATE FEEDBACK

Jingnong Yang and Douglas B. Williams

Center for Signal and Image Processing
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0250, USA

ABSTRACT

This paper describes a low-rate feedback algorithm for conveying partial channel state information—specifically, the dominant row subspace of the channel matrix—from the receiver to the transmitter in a continuously time-varying multiple-antenna environment. Since subspaces are points on a complex Grassmann manifold, variations in subspaces are treated as a piecewise geodesic process on the manifold. The receiver feeds back one bit to indicate the preferred sign of a random velocity matrix of the geodesic. Numerical results show that the performance of the proposed algorithm is better than the Grassmannian subspace packing approach at low-to-medium Doppler frequency and always better than the previously proposed gradient sign feedback scheme.

1. INTRODUCTION

It is well known that *multiple-input multiple-output* (MIMO) systems provide capacity enhancement over single-antenna systems. Space-time coding is designed to approach the channel capacity of MIMO systems [1, and references therein]. Most work on space-time coding deals with the "blind" case, where no knowledge of the forward channel is available to the transmitter.

In a MIMO system, if the transmitter has perfect knowledge of the underlying *channel state information* (CSI), a higher channel capacity is achievable compared to blind transmission through power allocation to the right singular subspace of the channel matrix. When the reciprocity of wireless channels does not hold, such as in frequency-division duplex, perfect CSI at the transmitter requires a high-rate feedback channel, which may not be practical, particularly in fast time-varying environments. Thus, the identification and utilization of partial CSI at the transmitter are important issues.

The benefits of partial CSI at the transmitter and the design of optimal transmission schemes with partial CSI are shown in [2]. Techniques proposed for attaining partial CSI include subspace quantization using Grassmannian subspace packing [3] and subspace tracking through a gradient sign feedback [4]. The efforts of attaining and utilizing partial CSI at the transmitter have also been extended to frequency-selective channels [5,6]. However, we will focus on flat-fading channels in this paper.

In [4], each unit of feedback is a single bit indicating the preferred sign of a random perturbation of the current transmit weight subspace, while each codeword in [3] has to be encoded into several bits. Therefore, with a very low-rate feedback channel receivers in [3] usually need channel prediction, and it is the predicted channel that is quantized. The complexity of the quantization algorithm in [3] is also considerable when the codebook size is large.

Here, a new partial CSI acquisition algorithm is proposed. We consider the transmit subspaces as points in a complex Grassmann manifold. Variations in subspaces are treated as a piecewise geodesic process in the Grassmann manifold. A one-bit feedback is utilized to indicate the preferred sign of a random velocity matrix of the geodesic.

The problem setting and an introduction of the geometry of Grassmann manifolds are given in Sections 2 and 3, respectively. The algorithm is demonstrated in Section 4. Section 5 provides an analysis of algorithm convergence. Numerical examples are shown in Section 6, and Section 7 concludes the paper.

Notation: Bold upper (lower) letters denote matrices (column vectors); $(\cdot)^H$ denotes Hermitian transpose; $\|\cdot\|_F$ stands for the Frobenius norm of a matrix; $E\{\cdot\}$ stands for expectation; $\mathrm{tr}\{\cdot\}$ is the trace of a matrix; $\Re\{\cdot\}$ stands for the real part of complex entries, $\mathrm{sign}(\cdot)$ for the signum function; $(\boldsymbol{I}_n)$ $\boldsymbol{I}$ denotes an $(n\times n)$ identity matrix; $\mathbf{0}_{m\times n}$ is an $m\times n$ all-zero matrix; $\mathrm{diag}(\boldsymbol{x})$ stands for a diagonal matrix with $\boldsymbol{x}$ on its diagonal; $\mathbb{C}^n$ denotes the n-dimensional complex space. The notation $\mathcal{CN}(\mu,\sigma^2)$ denotes the complex Gaussian distribution with mean μ and variance σ^2.

2. PROBLEM SETTING

We consider a flat-fading MIMO channel with N_r receive antennas and N_t transmit antennas, characterized by the following discrete-time input-output relationship,

$$\boldsymbol{y}[n] = \boldsymbol{H}[n]\boldsymbol{x}[n] + \boldsymbol{n}[n], \tag{1}$$

where $\boldsymbol{H}[n]$ is an $N_r\times N_t$ complex channel transfer matrix, and $\boldsymbol{n}[n]$ is an $N_r\times 1$ zero-mean complex Gaussian noise vector with covariance matrix $N_0\boldsymbol{I}$. The *singular value decomposition* (SVD) of $\boldsymbol{H}[n]$ is defined as

$$\boldsymbol{H}[n] = \boldsymbol{U}[n]\boldsymbol{\Lambda}[n]\boldsymbol{V}^H[n], \tag{2}$$

where $\boldsymbol{\Lambda}[n]$ has on its main diagonal the singular values of $\boldsymbol{H}[n]$ in descending order. The N_s-dimensional principal right-singular subspace of $\boldsymbol{H}[n]$ is spanned by the first N_s columns of $\boldsymbol{V}[n]$

$$\tilde{\boldsymbol{V}}[n] = \boldsymbol{V}(:,1:N_s)[n], \tag{3}$$

where we use MATLAB notation to denote a submatrix. Note that $\tilde{\boldsymbol{V}}[n]$ is an orthonormal matrix, i.e., $\tilde{\boldsymbol{V}}^H[n]\tilde{\boldsymbol{V}}[n] = \boldsymbol{I}_{N_s}$. The objective of the proposed algorithm is to track an $N_t\times N_s$ complex weight matrix $\boldsymbol{W}[n]$ with orthonormal columns that maps an $N_s\times$

1 complex vector $\boldsymbol{s}$ of coded message symbols to the transmitted signals

$$\boldsymbol{x}[n] = \boldsymbol{W}[n]\boldsymbol{s}[n]. \tag{4}$$

The tracking attempts to extract the principal right-singular subspaces, giving

$$\boldsymbol{W}[n]\boldsymbol{W}^H[n] = \tilde{\boldsymbol{V}}[n]\tilde{\boldsymbol{V}}^H[n]. \tag{5}$$

Such a $\boldsymbol{W}[n]$ is the orthonormal matrix that maximizes the received power [4], defined as

$$J[n] = \|\boldsymbol{H}[n]\boldsymbol{W}[n]\|_F^2. \tag{6}$$

It has been shown that perfect subspace tracking with equal power allocation among N_s dimensions has a power gain of precisely N_t/N_s over blind transmission [4].

Note that if $N_s > N_t/2$, we only need to track the orthogonal complement of $\tilde{\boldsymbol{V}}[n]$. Therefore, without loss of generality, we assume that $N_s \leq N_t/2$.

In the following section, we summarize the intrinsic properties of Grassmann manifolds.

3. GEODESICS OF GRASSMANN MANIFOLDS

In this section we summarize some properties of the geometry of Grassmann manifolds used in this paper [7]. A complex Grassmann manifold $\mathcal{G}_{N_t,N_s}$ contains all N_s-dimensional subspaces of $\mathbb{C}^{N_t}$. As mentioned previously, since the transmit beamforming matrix can be any orthonormal $N_t \times N_s$ matrix that forms an orthonormal basis of the principal right singular subspace of the MIMO channel, the Grassmann manifold is a natural description of the domain of the transmit beamforming matrices.

We introduce all necessary notations using time instant 0 as an example. The $N_t \times N_s$ orthonormal transmit weight matrix at time 0 is $\boldsymbol{W}[0]$. Denoted as $[\boldsymbol{W}[0]]$, a point in the Grassmann manifold is an equivalent class

$$[\boldsymbol{W}[0]] = \{\boldsymbol{W}[0]\boldsymbol{Q}_{N_s} : \boldsymbol{Q}_{N_s} \text{ is any } N_s \times N_s \text{ unitary matrix}\}, \tag{7}$$

i.e., a point in the Grassmann manifold is the set of all $N_t \times N_s$ orthonormal matrices whose columns span the same subspace as the columns of $\boldsymbol{W}[0]$. When performing computations on the Grassmann manifold, we use the matrix $\boldsymbol{W}[0]$ to represent the entire equivalence class. Let $\boldsymbol{Q}[0] = \begin{pmatrix} \boldsymbol{W}[0] & \boldsymbol{Z}[0] \end{pmatrix}$, where $\boldsymbol{Z}[0]$ contains as columns an orthonormal basis of the orthogonal complement of $[\boldsymbol{W}[0]]$. In other words, $\boldsymbol{Q}[0]$ is a unitary matrix.

A geodesic is the curve of shortest length between two points on a manifold. We will represent a stochastic process on $\mathcal{G}_{N_t,N_s}$ as a piecewise-geodesic curve with random velocities at individual pieces. Therefore, we need an explicit description of geodesics in $\mathcal{G}_{N_t,N_s}$. Geodesics in $\mathcal{G}_{N_t,N_s}$ starting from $\boldsymbol{W}[0]$ are parameterized by [7]

$$\boldsymbol{W}(t) = \boldsymbol{Q}[0]\exp(t\boldsymbol{B}[0])\boldsymbol{J}, \tag{8}$$

where

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{I}_{N_s} \\ \boldsymbol{0}_{(N_t-N_s)\times N_s} \end{pmatrix} \tag{9}$$

and is fixed throughout this paper. The matrix $\boldsymbol{B}[0]$ is further restricted to be of the form

$$\boldsymbol{B}[0] = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{A}^H[0] \\ \boldsymbol{A}[0] & \boldsymbol{0} \end{pmatrix}, \boldsymbol{A}[0] \in \mathbb{C}^{(N_t-N_s)\times N_s}. \tag{10}$$

We denote the point reached by the geodesic at time $t = 1$ as $\boldsymbol{W}[1]$, therefore $\boldsymbol{W}[1] = \boldsymbol{W}(1) = \boldsymbol{Q}[0]\exp(\boldsymbol{B}[0])\boldsymbol{J}$. The matrix $\boldsymbol{A}[0]$ determines the point $\boldsymbol{W}[1]$ uniquely given $\boldsymbol{W}[0]$. $\boldsymbol{A}[0]$ can be deemed as the velocity that takes $\boldsymbol{W}[0]$ to $\boldsymbol{W}[1]$ in unit time. In the MIMO transmit subspace tracking context, $[\boldsymbol{W}[0]]$ is the outdated knowledge of the transmit subspace at the transmitter, and $[\boldsymbol{W}[1]]$ is the current transmit subspace which we want to be as close to $\left[\tilde{\boldsymbol{V}}[1]\right]$ as possible. The essence of the proposed algorithm is to approximate $\boldsymbol{A}[0]$ using a random matrix together with a one-bit sign obtained from feedback.

We now summarize how to efficiently compute $\boldsymbol{W}[1]$ given $\boldsymbol{W}[0]$ and $\boldsymbol{A}[0]$, without performing matrix exponentials [8]. Let

$$\boldsymbol{A}[0] = \tilde{\boldsymbol{U}}_2\boldsymbol{\Theta}\boldsymbol{U}_1^H \tag{11}$$

be the compact SVD of $\boldsymbol{A}[0]$, where

$$\boldsymbol{\Theta} = \mathrm{diag}\begin{pmatrix} \theta_1 & \theta_2 & \ldots & \theta_{N_s} \end{pmatrix}. \tag{12}$$

The θ_k's, $k = 1, 2, \ldots, N_s$, are principal angles between the subspaces $[\boldsymbol{W}[0]]$ and $[\boldsymbol{W}[1]]$. Then it can be shown that

$$\boldsymbol{W}[1] = \boldsymbol{Q}[0]\begin{pmatrix} \boldsymbol{U}_1\boldsymbol{C} \\ \tilde{\boldsymbol{U}}_2\boldsymbol{S} \end{pmatrix}, \tag{13}$$

where $\boldsymbol{C}$ is a diagonal matrix with elements $\cos\theta_k$, $1 \leq k \leq N_s$, on the diagonal, and $\boldsymbol{S}$ is diagonal with elements $\sin\theta_k$. The computational complexity is $O(N_tN_s^2)$, far below the $O(N_t^3)$ implied by the expression $\exp(\boldsymbol{B}[0])$.

4. ALGORITHM DESCRIPTION

Before elaboration of the algorithm, we list our assumptions. In this paper, we assume that the time variation process is *independent and identically distributed* (i.i.d.) on each entry of the channel matrix $\boldsymbol{H}[n]$. From symmetry, we conjecture that entries of $\boldsymbol{A}[n]$ are i.i.d. $\mathcal{CN}(0, a^2)$, where a is a parameter depending on the Doppler frequency.

We assume that channel estimation at the receiver, feedback, and computation of the transmit weight matrix at the transmitter do not consume time and happen instantly. Similarly, for the Grassmannian subspace packing scheme with a codebook of size 2^N, we assume that the channel at time $n + N - 1$ is predicted by the receiver at time n, and the N-bit codeword is fed back at times $n, n + 1, \ldots, n + N - 1$. The dequantized beamforming matrix for time $n - 1$ is held constant at the transmitter for times $n - 1, n, \ldots, n + N - 2$.

For each feedback period, the random velocity $\boldsymbol{A}[n]$ is generated with i.i.d. $\mathcal{CN}(0, a^2)$ entries. We assume that the transmitter and the receiver have common knowledge of the value of $\boldsymbol{A}[n]$ at any time instant. $\boldsymbol{A}[n]$ can be conveyed from the transmitter to the receiver multiplexed with data symbols [4], or synchronously generated by pseudo random number generators at the transmitter and the receiver.

The feedback decision selects which sign-direction is preferable in terms of maximizing received power, as the result of an advancement along the geodesic. The received power at discrete time n is captured by the cost function $J[n]$ in (6). For each instance of the random matrix $\boldsymbol{A}[n]$, we approximate the new transmit subspace as one point in the Grassmann manifold reached by a geodesic in unit time, starting from the current transmit subspace, using $\boldsymbol{A}[n]$ as the velocity matrix. Using the parameterized

Table 1. Tracking algorithm summary

Initialize:	
$\quad \boldsymbol{W}[0] = \begin{pmatrix} \boldsymbol{I}_{N_s} \\ \boldsymbol{0} \end{pmatrix}$	(TX & RX)
for $n = 0 : \infty$	
$\quad \boldsymbol{A}[n]$=realization of Gaussian	(TX & RX)
$\quad$ Form unitary matrix $\boldsymbol{Q}[n] = [\boldsymbol{W}[n]\ \boldsymbol{Z}[n]]$	(TX & RX)
$\quad$ Compute $s[n]$ using (14)	(RX)
$\quad \boldsymbol{W}[n+1] = \boldsymbol{Q}[n]\exp(s[n]\boldsymbol{B}[n])\boldsymbol{J}$	(TX & RX)

geodesic in (8), with one bit of ambiguity, the new transmit subspace $\boldsymbol{W}[n+1]$ can take either value of $\boldsymbol{Q}[n]\exp(\pm\boldsymbol{B}[n])\boldsymbol{J}$. The binary feedback s is determined as

$$s[n] = \operatorname{sign}(\|\boldsymbol{H}[n+1]\boldsymbol{Q}[n]\exp(\boldsymbol{B}[n])\boldsymbol{J}\|_F^2 - \|\boldsymbol{H}[n+1]\boldsymbol{Q}[n]\exp(-\boldsymbol{B}[n])\boldsymbol{J}\|_F^2). \tag{14}$$

The weight matrix update at the transmitter is given by $\boldsymbol{W}[n+1] = \boldsymbol{Q}[n]\exp(s[n]\boldsymbol{B}[n])\boldsymbol{J}$. Since $\boldsymbol{W}[n+1] \in \mathcal{G}_{N_t,N_s}$, $\boldsymbol{W}[n+1]$ serves as the new orthonormal transmit matrix directly and no further orthonormalization is necessary, as opposed to the algorithm in [4], where a Gram-Schmidt QR factorization is required. The tracking algorithm is summarized in Table 1.

The parameter a controls the length of the geodesic arc. To see this, first introduce a matrix $\boldsymbol{A}_w[0]$ with i.i.d. $\mathcal{CN}(0,1)$ entries such that $\boldsymbol{A}[0] = a\boldsymbol{A}_w[0]$, for $a > 0$. The arc length distance between $\boldsymbol{W}[0]$ and $\boldsymbol{W}[1]$ is defined as [7] $d(\boldsymbol{W}[0], \boldsymbol{W}[1]) = \left(\sum_{i=1}^{N_s} \theta_i^2\right)^{1/2}$, where the θ_i's are defined in (12). Using (11), it can be shown that $d(\boldsymbol{W}[0], \boldsymbol{W}[1]) = \|\boldsymbol{A}[0]\|_F = a\|\boldsymbol{A}_w[0]\|_F$, i.e., the parameter a is proportional to the average arc length of the geodesic. Intuitively, the arc length of the geodesic traversed in unit time for high Doppler frequencies is larger than for low Doppler frequencies. Therefore, the parameter a should be chosen monotonically with Doppler frequency.

5. GRADIENT EXTRACTION

In this section, we analyze the convergence behavior assuming that $\boldsymbol{H}$ is static and non-random. We first simplify the expression of s in (14). When the channel is slowly varying, the norm of $\boldsymbol{A}[0]$ is small. Equivalently, the values of the θ_ks are small. In this case, a first-order approximation gives $\boldsymbol{C} \approx \boldsymbol{I}$ and $\boldsymbol{S} \approx \boldsymbol{\Theta}$. Using (10), (11), (13), and the above approximation, we have

$$s[0] \approx \operatorname{sign}\operatorname{tr}\Re(\boldsymbol{W}^H[0]\boldsymbol{H}^H[0]\boldsymbol{H}[0]\boldsymbol{Z}[0]\boldsymbol{A}[0]). \tag{15}$$

Differentiation of $\boldsymbol{W}(t)$ in (8) gives the direction of the random geodesic with one-bit feedback

$$\dot{\boldsymbol{W}}(0) = \dot{\boldsymbol{W}}(t)|_{t=0} = s[0]\boldsymbol{Q}[0]\boldsymbol{B}[0]\boldsymbol{J} = s[0]\boldsymbol{Z}[0]\boldsymbol{A}[0]. \tag{16}$$

The expectation of this direction is $E\{\dot{\boldsymbol{W}}(0)\}$=$\boldsymbol{Z}[0]E\{s[0]\boldsymbol{A}[0]\}$. With (15) and the results of [4, Appendix A], we have

$$E\{\dot{\boldsymbol{W}}(0)\} = a\sqrt{\frac{1}{\pi}}\frac{\boldsymbol{Z}[0]\boldsymbol{Z}^H[0]\boldsymbol{H}^H[0]\boldsymbol{H}[0]\boldsymbol{W}[0]}{\|\boldsymbol{Z}^H[0]\boldsymbol{H}^H[0]\boldsymbol{H}[0]\boldsymbol{W}[0]\|_F}. \tag{17}$$

At the point $\boldsymbol{W}[0]$, the gradient of the cost function J defined on the Grassmann manifold is given by [7]

$$2\boldsymbol{Z}[0]\boldsymbol{Z}^H[0]\boldsymbol{H}^H[0]\boldsymbol{H}[0]\boldsymbol{W}[0]. \tag{18}$$

Comparing with (17), we see immediately that the expected direction of the geodesic is proportional to the gradient. Among all tangential directions, this gradient points in the direction of maximum increase of the cost function. In contrast, the expected direction of subspace updating in [4] is proportional to the gradient in the Euclidean sense.

To consider the estimation error, we define the error matrix $\boldsymbol{E} = \dot{\boldsymbol{W}}(0) - E\{\dot{\boldsymbol{W}}(0)\}$. One measurement of the estimation error is related to the second moment of the error matrix through

$$\operatorname{tr}(E\{\boldsymbol{E}^H\boldsymbol{E}\}) = a^2\left((N_t - N_s)N_s - 1/\pi\right). \tag{19}$$

While in [4], the estimation error embodied by a similarly defined error matrix $\boldsymbol{E}$ [4, Eq. (60)] is given by

$$\operatorname{tr}(E\{\boldsymbol{E}^H\boldsymbol{E}\}) = 2\beta^2\left(N_tN_s - 1/\pi\right), \tag{20}$$

where β is an adaptation parameter controlling the step size. Ignoring the parameters a, β and the constant two coming from the different variance of the entries of random matrices, it can be seen that the estimation error of our algorithm is smaller. This advantage arises because the number of degrees of freedom of $\boldsymbol{A}[n]$ is only $(N_t - N_s)N_s$, which is less than N_tN_s, the dimension of random perturbation in [4].

6. NUMERICAL RESULTS

A Monte Carlo simulation was performed to test the performance of the proposed transmit subspace tracking algorithm in the setting of $N_s = N_r = 2$ and $N_t = 8$. The receiver channel estimation error was modeled as AWGN added to the true channel matrix. The channel estimation error was 20 dB below the channel power. The feedback channel was free of error. The channel model was independent Rayleigh flat fading with time correlation generated by Jakes' method. The relationship between the Doppler frequency F_D and the feedback rate F_{FB} was captured by the ratio F_{FB}/F_D. All the numerical results were compared with the feedback schemes in [3, 4]. For [3], the size of the Grassmannian subspace packing codebook was $2^N = 1024$, and the receiver utilized a linear predictor of length 26 or 50 to predict the channel matrix $\boldsymbol{H}$ nine intervals ahead. Optimal values of a and β [4] versus Doppler frequency were determined through numerical search to minimize the mean estimation error. In the figure legends, *ideal* denotes ideal subspace tracking, *geodesic* denotes our algorithm, *gradient sign* stands for [4], *Grassmann 26 (50)* stands for [3] with linear predictor length 26 (50), and *blind* stands for blind transmission.

The average performance of tracking constant channels is illustrated in Fig. 1. The result is in dB relative to the cost of ideal subspace tracking. As can be seen, larger values of a and β resulted in faster convergence but larger stable-state error. Our algorithm showed faster convergence and smaller stable-state error than the gradient sign algorithm.

The ergodic capacities achieved by different algorithms are given in Fig. 2 for $F_{FB}/F_D = 1000$. Our algorithm shows a 2-3 dB gain over the gradient sign algorithm across the medium-to-high SNR range. For this given value of F_{FB}/F_D, both our algorithm and the gradient sign feedback outperform the Grassmannian subspace packing algorithm.

Finally, we compare the tracking performance in terms of cost function J. The feedback rate F_{FB} is fixed at 6000 Hz, and the

Doppler frequency varies from 1 Hz to 100 Hz. The result is normalized to the performance of ideal subspace tracking (J_{opt}) and is shown in Fig. 3. Our algorithm performs uniformly better than the gradient sign feedback, and is better than the Grassmannian subspace packing algorithm for F_D as high as 55 Hz. When the feedback rate is low or the Doppler frequency is high, the tracking algorithms' performance is limited, and the batch feedback algorithm performs better.

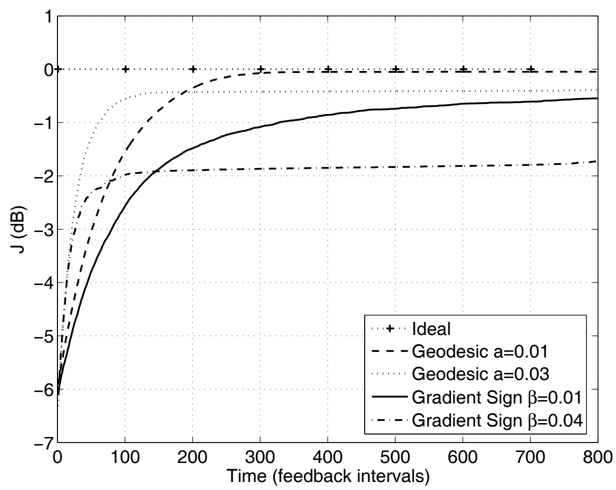

Fig. 1. Convergence transient of received power.

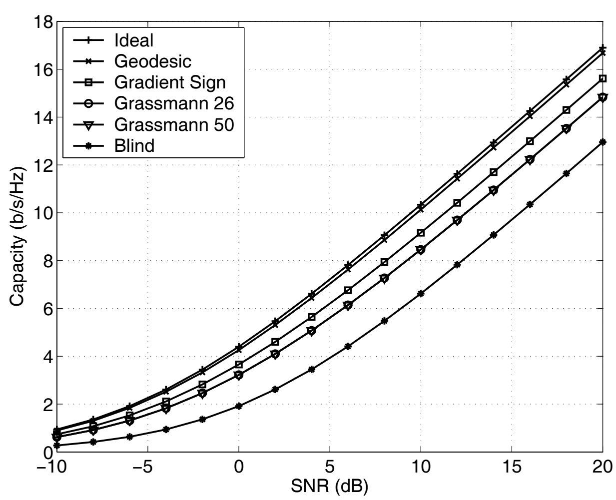

Fig. 2. Ergodic capacity versus SNR at $F_{FB}/F_D = 1000$.

7. CONCLUSION

In the proposed transmission subspace tracking algorithm, variations in transmit weight matrix are treated as a piecewise geodesic process in Grassmann manifolds. The expected direction of the geodesic at the starting point is proportional to the gradient of a cost function in the Riemannian sense. The proposed algorithm tries to track fewer parameters than the gradient sign algorithm

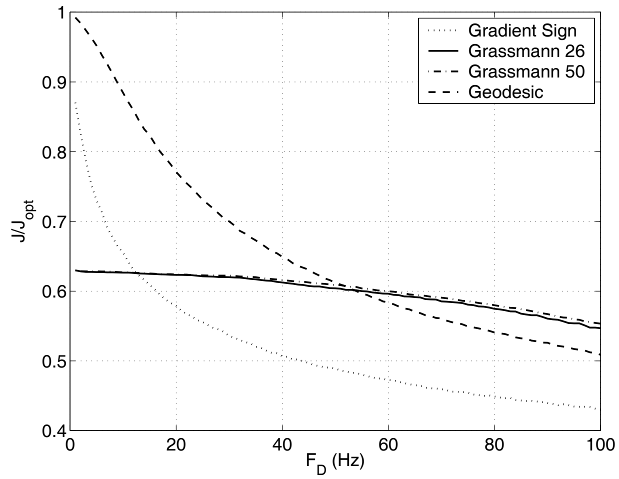

Fig. 3. Mean cost function versus Doppler frequency.

and achieves better performance. Numerical results show that the performance of the adaptive algorithm approaches ideal subspace tracking for feedback rates on the order of 1000 times the channel Doppler frequency. Compared with a Grassmannian subspace packing quantization algorithm, our algorithm has better performance at low-to-medium Doppler frequency and does not incur the complexity of quantization and long-range channel prediction.

8. REFERENCES

[1] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications - a key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[2] E. Visotsky and U. Madhow, "Space-time transmit precoding with imperfect feedback," *IEEE Trans. Inform. Theory*, vol. 47, no. 6, pp. 2632–2639, Sept. 2001.

[3] D. J. Love and R. W. Heath, Jr., "Limited feedback precoding for spatial multiplexing systems using linear receivers," in *Proc. IEEE MILCOM*, Boston, MA, Oct. 2003.

[4] B. C. Banister and J. R. Zeidler, "Feedback assisted transmission subspace tracking for MIMO systems," *IEEE J. Selected Areas Commun.*, vol. 21, no. 3, pp. 452–463, Apr. 2003.

[5] J. Yang and Y. G. Li, "Low complexity OFDM MIMO system based on channel correlations," in *Proc. IEEE Global Telecommun. Conf.*, San Francisco, CA, Dec. 2003, vol. 2, pp. 591–595.

[6] J. Choi and R. W. Heath, Jr., "Interpolation based transmit beamforming for MIMO-OFDM with limited feedback," in *Proc. IEEE Int. Conf. Commun.*, June 2004, vol. 1, pp. 249–253.

[7] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix. Anal. Appl.*, vol. 20, no. 2, pp. 303–353, Oct. 1998.

[8] K. A. Gallivan, A. Srivastava, X. Liu, and P. V. Dooren, "Efficient algorithms for inferences on Grassmann manifolds," in *Proc. IEEE Workshop on Statistical Signal Processing*, Sept. 2003, pp. 315–318.