On the Quantization and Prediction for Precoded MIMO with Delayed Limited Feedback

Dalin Zhu and Ming Lei

The Department of Wireless Communications
NEC Laboratories China (NLC)
11F Bldg.A, Innovation Plaza TusPark, Beijing, China 100084
E-mail: {zhu_dalin, lei_ming}@nec.cn

***Abstract*—In closed-loop precoded multiple-input multiple-output (MIMO) systems, the representative quantized channel state information (CSI) usually becomes outdated before its actual use at the transmitter due to the delayed feedback. In this paper, by exploiting the differential geometric properties of the Grassmannian manifold, we invoke the use of Grassmannian subspace predictive coding (GSPC) to compensate for this feedback delay. More specifically, a novel limited feedback strategy is presented which consists of a two-stage optimization process at the receiver. The first-stage optimization is accomplished by the quantization, which is custom designed as a function of the optimally predicted CSI under the framework of GSPC; after the first-stage codeword selection, the second-stage optimization is performed by optimizing the prediction on the Grassmannian manifold. We show by simulations that, the proposed encoding/decoding algorithm exhibits promised system performance over the existing prediction methods, for both spatial multiplexing and transmit beamforming MIMO, under the same amount of feedback bits.**

I. Introduction

Multiple-input multiple-output (MIMO) system is capable of supporting high throughput and highly reliable wireless transmissions through multiplexing gain and diversity gain, respectively [1], [2]. Transmit precoding is a promising technique in MIMO to achieve higher channel capacity. However, transmit precoding requires the channel state information (CSI) available at the transmitter in closed-loop MIMO systems. Limited feedback is commonly used to convey the CSI to the transmitter. At the receiver, by quantizing the estimated CSI using a fixed off-line designed codebook, only the index of the selected codeword is fed back to the transmitter for CSI's reconstruction [3]. Quantization on the Grassmannian manifold can be referred to [4]. In most prior works, one-shot memoryless limited feedback strategy is performed by adopting the block fading channel model [5][6][7]. However, in practice, due to the mobility in the propagation environment, the wireless channels usually exhibit memories, which can be characterized by the temporal correlations [8]. In addition, due to the feedback delay, the representative quantized CSI usually becomes outdated prior to its actual use at the transmitter. This feedback delay is brought by the channel-access protocols overhead and/or signal processing intervals, which would significantly degrade the system performance.

Numerous research efforts have devoted to designing efficient feedback strategies for time-selective MIMO channels with memories [8]-[12]. However, none of these efforts concentrates on accurately tracking the future CSI to compensate for the feedback delay. It has been well recognized that *prediction* is a promising technique to overcome this problem. Authors in [13] and [14] propose to apply linear predictive coding to provide high resolution CSI. In [15] and [16], based on a first-order autoregressive (AR1) dynamic fading model, predictive vector quantization (PVQ) is employed to track the forthcoming CSI. Grassmannian subspace tracking (GST) is deployed in [17] to predict the CSI in the case of spatial multiplexing MIMO. However, the proposed GST algorithm highly relies on the propagation model, which would be of critical concern regarding practical implementations. Grassmannian predictive coding (GPC) algorithm is investigated in [18] for transmit beamforming MIMO. By exploiting the differential geometric properties of the Grassmannian manifold, a manifold constrained prediction framework is developed. However, in [18], as both the direction and the amplitude of the error tangent vector are separately quantized, the overall quantization resolution may not be ensured especially at relatively low feedback rate. Additionally, the extension of the proposed GPC algorithm to the spatial multiplexing case is not straightforward.

In this work, a new CSI tracking scheme is proposed and analyzed for closed-loop precoded MIMO assuming delayed feedback. As indicated in [6], the transmit subspaces are points on a complex Grassmannian manifold. In temporally correlated MIMO channels, these points are related by their corresponding tangents [19]. Therefore, by exploiting these features of the Grassmannian manifold, we propose to apply a novel prediction framework, named Grassmannian subspace predictive coding (GSPC) to solve the feedback delay problem without constraint on the transmission rank. The associated limited feedback scheme consists of a two-stage optimization process at the receiver. The first-stage optimization is accomplished by the quantization, which is custom designed as a function of the optimally predicted CSI. After the codeword selection, the second-stage optimization is performed by maximizing the system throughput. Additionally, we show that

978-1-4673-0990-5/12/$31.00 ©2012 IEEE

the prediction framework investigated in [18] is actually a special case of our proposed GSPC for rank-1 transmission (i.e., transmit beamforming). Numerical results show that our proposed scheme outperforms the existing prediction methods for both spatial multiplexing and transmit beamforming MIMO, in terms of the throughput and error rate performances, under the same amount of feedback bits.

The rest of this paper is organized as follows: in section II, the general framework of GSPC is studied; then, the proposed encoding/decoding algorithm is presented in section III; in section IV, we use the transmit beamforming as an example to illustrate and analyze; numerical results are provided in section V; finally, we conclude our paper in section VI.

II. Grassmannian Subspace Predictive Coding

We start by briefly reviewing the conventional one-shot memoryless feedback strategy for block fading MIMO channel. At time k, the channel matrix $\boldsymbol{H}[k]$ is assumed to be a $M_r \times M_t$ block matrix with each entry distributed according to $\mathcal{CN}(0,1)$. The M-dimensional principal right singular matrix $\widetilde{\boldsymbol{U}}[k]$ can be obtained as the first M columns of matrix $\boldsymbol{U}[k]$, which is the $M_t \times M_t$ right singular matrix of the singular value decomposition (SVD) of $\boldsymbol{H}[k]$. Assume that $\mathcal{W} = \{\boldsymbol{W}_1, \cdots, \boldsymbol{W}_L\}$ is a fixed off-line designed codebook, where $\boldsymbol{W}_i \in \boldsymbol{U}^{M_t \times M}$, the set of $M_t \times M$ orthonormal complex matrices, and L is the total number of codewords. Given a codeword $\boldsymbol{W}_i$ and the channel realization $\boldsymbol{H}[k]$, the instantaneous channel capacity of the unitary precoded MIMO is calculated as

$$I\left(\boldsymbol{H}[k], \boldsymbol{W}_i\right) = \log\det\left(\boldsymbol{I}_{M_r,M_r} + \frac{E_s}{N_0}\boldsymbol{H}[k]\boldsymbol{W}_i\boldsymbol{W}_i^H\boldsymbol{H}^H[k]\right). \quad (1)$$

Here, E_s is the total transmit power per symbol from all M_t transmit antennas; N_0 denotes the noise variance. As our primary goal is to enhance the system throughput, the quantization criterion is therefore set to be [6]

$$\boldsymbol{W}[k] = \arg\max_{\boldsymbol{W}_i \in \mathcal{W}} I\left(\boldsymbol{H}[k], \boldsymbol{W}_i\right), \quad (2)$$

where $\boldsymbol{W}[k]$ is the selected codeword at time k and its index in $\mathcal{W}$ is fed back to the transmitter for CSI's reconstruction.

However, in practice, (i) the wireless channels exhibit memories; and (ii) the representative quantized CSI usually becomes outdated prior to its application at the transmitter. As the transmit subspaces are points on a complex Grassmannian manifold, (i) implies that the adjacent transmit subspaces are related by the channel variations as along a geodesic, which is the curve of the shortest length between two points on a manifold. By exploiting this, we invoke the use of GSPC to solve (ii). The framework of GSPC is depicted in the following. The geodesic that emanates from $\widetilde{\boldsymbol{U}}[k-1]$ to $\widetilde{\boldsymbol{U}}[k]$ is parameterized as [19]

$$\widetilde{\boldsymbol{U}}[k] = \boldsymbol{Z}[k-1]\exp\left(\boldsymbol{B}[k-1]\right)\boldsymbol{I}_{M_t,M}. \quad (3)$$

Here, $\boldsymbol{Z}[k-1] = \left(\widetilde{\boldsymbol{U}}[k-1] \;\; \widetilde{\boldsymbol{U}}_\perp[k-1]\right)$, where $\widetilde{\boldsymbol{U}}_\perp[k-1]$ is the orthogonal complement of $\widetilde{\boldsymbol{U}}[k-1]$. Straightforwardly, $\boldsymbol{Z}[k-1]$ is a square unitary matrix that has the dimension of $M_t \times M_t$. $\boldsymbol{B}[k-1]$ is a skew Hermitian matrix[1], expressed as

$$\boldsymbol{B}[k-1] = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{A}^H[k-1] \\ \boldsymbol{A}[k-1] & \boldsymbol{0} \end{pmatrix}, \quad (4)$$

where $\boldsymbol{A}[k-1] \in \boldsymbol{C}^{M_t-M,M}$, the set of $(M_t-M)\times M$ complex matrices. (3) and (4) imply that by given $\widetilde{\boldsymbol{U}}[k-1]$ and $\widetilde{\boldsymbol{U}}[k]$, $\boldsymbol{A}[k-1]$ can be uniquely determined. Hence, we express $\boldsymbol{A}[k-1]$ in terms of $\widetilde{\boldsymbol{U}}[k-1]$ and $\widetilde{\boldsymbol{U}}[k]$ by performing the cosine-sine decomposition, given as

$$\begin{pmatrix} \widetilde{\boldsymbol{U}}^H[k-1]\widetilde{\boldsymbol{U}}[k] \\ \widetilde{\boldsymbol{U}}_\perp^H[k-1]\widetilde{\boldsymbol{U}}[k] \end{pmatrix} = \begin{pmatrix} \boldsymbol{U}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{U}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{C} \\ \boldsymbol{S} \end{pmatrix} \boldsymbol{V}_1^H, \quad (5)$$

where $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ are $M\times M$ and $(M_t-M)\times M$ orthonormal matrices, respectively; $\boldsymbol{C}$ ($\boldsymbol{S}$) is a diagonal matrix with elements $\cos\theta_j$'s ($\sin\theta_j$'s) on its diagonal; and θ_j's ($1 \le j \le M$) are the principal angels between $\widetilde{\boldsymbol{U}}[k-1]$ and $\widetilde{\boldsymbol{U}}[k]$ [19]; $\boldsymbol{V}_1$ is the right singular matrix of the SVD of $\widetilde{\boldsymbol{U}}^H[k-1]\widetilde{\boldsymbol{U}}[k]$. Then, one possible choice for the velocity matrix $\boldsymbol{A}[k-1]$ turns out to be [17]

$$\boldsymbol{A}[k-1] = \boldsymbol{U}_2\boldsymbol{\Theta}\boldsymbol{U}_1^H, \quad (6)$$

where $\boldsymbol{\Theta} = \mathrm{diag}(\theta_1, \cdots, \theta_M)$. It is worth noting here that besides the velocity matrix $\boldsymbol{A}[k-1]$, the tangent between two adjacent points on a complex Grassmannian manifold can also be used to characterize their relations. Here, the tangent $\boldsymbol{\Delta}[k-1]$ that starts from $\widetilde{\boldsymbol{U}}[k-1]$ to $\widetilde{\boldsymbol{U}}[k]$ is calculated as [19]

$$\boldsymbol{\Delta}[k-1] = \widetilde{\boldsymbol{U}}_\perp[k-1]\boldsymbol{A}[k-1]. \quad (7)$$

Equipped with (7), by exploring the smooth structure of the Grassmannian manifold, we parallel transport[2] $\boldsymbol{\Delta}[k-1]$ along the geodesic with the new base point $\widetilde{\boldsymbol{U}}[k]$, obtaining

$$\widetilde{\boldsymbol{\Delta}}[k] = \left(\boldsymbol{I} - \widetilde{\boldsymbol{U}}[k-1]\boldsymbol{U}_d\sin\boldsymbol{\Sigma}_d t + \boldsymbol{V}_d\cos\boldsymbol{\Sigma}_d t\right)\boldsymbol{\Delta}[k-1], \quad (8)$$

where $\boldsymbol{V}_d\boldsymbol{\Sigma}_d\boldsymbol{U}_d^H$ is the compact SVD of $\boldsymbol{\Delta}[k-1]$; t is the step size parameter to be optimized and the identity matrix has the dimension of $M_t \times M$. Based on this, the predicted transmit precoding matrix along the geodesic from $\widetilde{\boldsymbol{U}}[k-1]$ to $\widetilde{\boldsymbol{U}}[k]$ can be derived as

$$\overline{\boldsymbol{U}}[k+1] = \widetilde{\boldsymbol{U}}[k]\boldsymbol{U}_p\cos\boldsymbol{\Sigma}_p t + \boldsymbol{V}_p\sin\boldsymbol{\Sigma}_p t, \quad (9)$$

where $\boldsymbol{V}_p\boldsymbol{\Sigma}_p\boldsymbol{U}_p^H$ is the compact SVD of $\widetilde{\boldsymbol{\Delta}}[k]$. After all, the step size parameter is optimized by maximizing the system throughput, given as

$$t_{\mathrm{opt}} = \arg\max_{t\in[0,1]} E\left[I\left(\boldsymbol{H}[k+1], \overline{\boldsymbol{U}}[k+1]\right)\right], \quad (10)$$

where $E(\cdot)$ denotes the expectation over the channel realizations. Then, by substituting (10) into (9), the forthcoming CSI can be predicted as

$$\widehat{\boldsymbol{U}}[k+1] = \widetilde{\boldsymbol{U}}[k]\boldsymbol{U}_p\cos\boldsymbol{\Sigma}_p t_{\mathrm{opt}} + \boldsymbol{V}_p\sin\boldsymbol{\Sigma}_p t_{\mathrm{opt}}. \quad (11)$$

[1]The explicit definition of $\boldsymbol{B}[k]$ can be addressed to [19]. The calculation of the matrix exponential of $\boldsymbol{B}[k]$ can be referred to eqn.(15) in [17].

[2]The explicit definition of the parallel transport can be addressed to [19].

III. Quantization and Prediction

By far, it becomes clear that we should somehow let the transmitter be aware of the optimally predicted CSI $\widehat{\boldsymbol{U}}[k+1]$ by means of feedback. Under the framework of GSPC, we propose a novel two-stage optimization process at the receiver. The first-stage optimization is accomplished by the quantization, which is custom determined as a function of the optimally predicted CSI. After the first-stage codeword selection, the second-stage optimization is achieved by optimizing the prediction on the Grassmannian manifold. Outcomes from the two stages are correspondingly quantized and fed back to the transmitter in terms of a limited number of bits.

A. First-stage quantization

At time k, for a given codeword $\boldsymbol{W}_i$, the corresponding predicted precoding matrix is computed as

$$\widetilde{\boldsymbol{W}}_i = \boldsymbol{W}_i \boldsymbol{Q}_i \cos \boldsymbol{S}_i t_i + \boldsymbol{P}_i \sin \boldsymbol{S}_i t_i, \tag{12}$$

where $\boldsymbol{P}_i \boldsymbol{S}_i \boldsymbol{Q}_i^H$ is the compact SVD of $\widetilde{\boldsymbol{\Delta}}_i$, which is the parallel transport tangent along the geodesic from $\boldsymbol{W}[k-1]$ to $\boldsymbol{W}_i$; $\boldsymbol{W}[k-1]$ is the quantized CSI at time $k-1$; t_i denotes the associated step size parameter. The calculation of $\widetilde{\boldsymbol{\Delta}}_i$ is similar to (8) by first computing the tangent from $\boldsymbol{W}[k-1]$ to $\boldsymbol{W}_i$. t_i's are first initialized as t_{opt}, and then configured as the step size parameter optimized at the second stage. Under the framework of GSPC, the optimally predicted CSI has been derived in (11). Therefore, for a given $\widehat{\boldsymbol{U}}[k+1]$, the new quantization criterion becomes to

$$\boldsymbol{W}[k] = \arg\max\nolimits_{\boldsymbol{W}_i \in \mathcal{W}} \left\| \widehat{\boldsymbol{U}}^H[k+1] \widetilde{\boldsymbol{W}}_i \right\|_F^2, \tag{13}$$

which is a function of the optimal prediction that adequately accounts for the feedback delay. Hence, the codebook can be simply constructed by using the minimum distance based metric [6], without extra design efforts. Additionally, we notice that if only one fixed off-line designed codebook is deployed, the same codeword may possibly be selected for both $\boldsymbol{W}[k-1]$ and $\boldsymbol{W}[k]$. In this case, the corresponding error tangent is calculated as null, which would terminate the prediction process. Therefore, we propose to employ two different[3] codebooks to alternatively quantize the previous and current CSIs. For instance, two different codebooks are first constructed by using, e.g., Lloyds algorithm [20], and are denoted by $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively. If $\boldsymbol{W}[k-1]$ is chosen from, say, $\mathcal{W}_1$, then $\boldsymbol{W}[k]$ should be selected from $\mathcal{W}_2$. By performing this, the previous and current quantized CSIs $\boldsymbol{W}[k-1]$ and $\boldsymbol{W}[k]$ will never be the same without loss of quantization resolution. In the following section, some mathematical interpretations are provided on the proposed quantization approach by using the transmit beamforming as an example to illustrate.

B. Second-stage prediction optimization

At the first-stage quantization, $\boldsymbol{W}[k]$ that maximizes $\left\| \widehat{\boldsymbol{U}}^H[k+1] \widetilde{\boldsymbol{W}}_i \right\|_F^2$ is chosen and its index in $\mathcal{W}$ is fed back to the transmitter. Therefore, at time k, the predicted CSI along the geodesic direction from $\boldsymbol{W}[k-1]$ to $\boldsymbol{W}[k]$ can be calculated as

$$\widetilde{\boldsymbol{W}}[k+1] = \boldsymbol{W}[k]\boldsymbol{Q}[k] \cos \boldsymbol{S}[k] \tilde{t}_{\text{opt}} + \boldsymbol{P}[k] \sin \boldsymbol{S}[k] \tilde{t}_{\text{opt}}. \tag{14}$$

Similarly, $\boldsymbol{P}[k]\boldsymbol{S}[k]\boldsymbol{Q}^H[k]$ is the compact SVD of $\overline{\boldsymbol{\Delta}}[k]$, the parallel transport tangent along the geodesic from $\boldsymbol{W}[k-1]$ to $\boldsymbol{W}[k]$; $\tilde{t}_{\text{opt}}$ is the optimized step size parameter obtained by maximizing the system throughput (similar to (10)). We note here that the closed-form expression of $\tilde{t}_{\text{opt}}$ is intractable and numerical search has to be conducted. As a long-term feedback, $\tilde{t}_{\text{opt}}$ is infrequently sent back to the transmitter to assist CSI's reconstruction.

IV. Transmit Beamforming

As a special case of precoding, transmit beamforming (i.e., $M = 1$) provides full diversity gain in multi-antenna systems. In this section, we first show that the prediction framework studied in [18] is actually a special case of our presented GSPC for rank-1 transmission. Then, we provide some insights on the proposed quantization criterion (i.e., (13)). We start by calculating the tangent between $\widetilde{\boldsymbol{U}}[k-1]$ and $\widetilde{\boldsymbol{U}}[k]$ assuming $M = 1$. According to (6) and (7), we have

$$\begin{aligned} \boldsymbol{\Delta}[k-1] &= \widetilde{\boldsymbol{U}}_\perp[k-1]\boldsymbol{A}[k-1] \\ &= \widetilde{\boldsymbol{U}}_\perp[k-1]\boldsymbol{U}_2 \boldsymbol{S} \frac{\boldsymbol{\Theta}}{\boldsymbol{SC}} \boldsymbol{C}^H \boldsymbol{U}_1^H. \end{aligned} \tag{15}$$

Based on (5), (15) can be rewritten as

$$\begin{aligned} \boldsymbol{\Delta}[k-1] &= \frac{\boldsymbol{\Theta}}{\boldsymbol{SC}} \widetilde{\boldsymbol{U}}_\perp[k-1] \widetilde{\boldsymbol{U}}_\perp^H[k-1] \widetilde{\boldsymbol{U}}[k] \widetilde{\boldsymbol{U}}^H[k] \widetilde{\boldsymbol{U}}[k-1] \\ &= \frac{\boldsymbol{\Theta}}{\boldsymbol{SC}} \widetilde{\boldsymbol{U}}^H[k] \widetilde{\boldsymbol{U}}[k-1] \\ &\times \left(\widetilde{\boldsymbol{U}}[k] - \widetilde{\boldsymbol{U}}[k-1] \widetilde{\boldsymbol{U}}^H[k-1] \widetilde{\boldsymbol{U}}[k] \right). \end{aligned} \tag{16}$$

Let $\rho = \widetilde{\boldsymbol{U}}^H[k-1]\widetilde{\boldsymbol{U}}[k]$ denote the inner product; straightforwardly, we have $\rho = \boldsymbol{C}$. Let $d = \sqrt{1-|\rho|^2}$ denote the chordal distance between $\widetilde{\boldsymbol{U}}[k-1]$ and $\widetilde{\boldsymbol{U}}[k]$ on $\mathcal{G}_{M_t,1}$; as the chordal distance can be interpreted as the sine of the principal angel between two subspaces, we have $d = \boldsymbol{S}$. By substituting the above results into (16), we obtain

$$\begin{aligned} \boldsymbol{\Delta}[k-1] &= \frac{\boldsymbol{\Theta}}{d\rho} \rho^* \left(\widetilde{\boldsymbol{U}}[k] - \rho \widetilde{\boldsymbol{U}}[k-1] \right) \\ &= \frac{\boldsymbol{\Theta}}{d\rho} \rho^* \rho \left(\widetilde{\boldsymbol{U}}[k]/\rho - \widetilde{\boldsymbol{U}}[k-1] \right) \\ &= \tan^{-1}\left(\frac{d}{|\rho|} \right) \frac{\widetilde{\boldsymbol{U}}[k]/\rho - \widetilde{\boldsymbol{U}}[k-1]}{d/\rho}. \end{aligned} \tag{17}$$

Next, we rewrite the parallel transport tangent $\widetilde{\boldsymbol{\Delta}}[k]$ for $M = 1$. The original expression of (8) is given as[4] (*Theorem*

[3] By difference, we mean that codewords in the two codebooks are different from each other.

[4] (8) is obtained by transforming the original expression by performing SVD operations and decompositions on the exponential map.

2.4 in [19])

$$\widetilde{\boldsymbol{\Delta}}[k] = \boldsymbol{Z}[k-1]\exp\left(\begin{array}{cc}\boldsymbol{0} & -\boldsymbol{A}^H[k-1]\\ \boldsymbol{A}^H[k-1] & \boldsymbol{0}\end{array}\right) \times \left(\begin{array}{c}\boldsymbol{0}\\ \boldsymbol{A}[k-1]\end{array}\right). \quad (18)$$

Denoting $\boldsymbol{U} = \mathrm{diag}\,(\boldsymbol{U}_1, \boldsymbol{U}_2)$, it is shown in [17] that,

$$\exp\left(\begin{array}{cc}\boldsymbol{0} & -\boldsymbol{A}^H[k-1]\\ \boldsymbol{A}^H[k-1] & \boldsymbol{0}\end{array}\right) = \boldsymbol{U}\boldsymbol{\Psi}\boldsymbol{U}^H, \quad (19)$$

where,

$$\boldsymbol{\Psi} = \left(\begin{array}{cc}\boldsymbol{C} & -\boldsymbol{S}\\ \boldsymbol{S} & \boldsymbol{C}^H\end{array}\right). \quad (20)$$

By plugging (6) and (19) into (18), we have

$$\begin{aligned}\widetilde{\boldsymbol{\Delta}}[k] &= -\boldsymbol{S}\boldsymbol{\Theta}\widetilde{\boldsymbol{U}}[k-1] + \boldsymbol{C}^H\widetilde{\boldsymbol{U}}_{\perp}[k-1]\boldsymbol{U}_2\boldsymbol{\Theta}\boldsymbol{U}_1^H\\ &= -d\tan^{-1}\left(\frac{d}{|\rho|}\right)\widetilde{\boldsymbol{U}}[k-1]\\ &+ \rho^*\tan^{-1}\left(\frac{d}{|\rho|}\right)\frac{\widetilde{\boldsymbol{U}}[k]/\rho - \widetilde{\boldsymbol{U}}[k-1]}{d/\rho}\\ &= \tan^{-1}\left(\frac{d}{|\rho|}\right)\frac{\rho^*\widetilde{\boldsymbol{U}}[k] - \widetilde{\boldsymbol{U}}[k-1]}{d}. \end{aligned} \quad (21)$$

Consecutively, as for $M = 1$, $\widetilde{\boldsymbol{\Delta}}[k]$ turns out to be a $M_t \times 1$ vector, we have $\boldsymbol{V}_p = \widetilde{\boldsymbol{\Delta}}[k]/\|\widetilde{\boldsymbol{\Delta}}[k]\|$, $\boldsymbol{\Sigma}_p = \|\widetilde{\boldsymbol{\Delta}}[k]\|$, and $\boldsymbol{U}_p = 1$. Hence, (9) can be reformulated as

$$\overline{\boldsymbol{U}}[k+1] = \widetilde{\boldsymbol{U}}[k]\cos\left(\|\widetilde{\boldsymbol{\Delta}}[k]\|t\right) + \frac{\widetilde{\boldsymbol{\Delta}}[k]}{\|\widetilde{\boldsymbol{\Delta}}[k]\|}\sin\left(\|\widetilde{\boldsymbol{\Delta}}[k]\|t\right), \quad (22)$$

and therefore,

$$\widehat{\boldsymbol{U}}[k+1] = \widetilde{\boldsymbol{U}}[k]\cos\left(\|\widetilde{\boldsymbol{\Delta}}[k]\|t_{\mathrm{opt}}\right) + \frac{\widetilde{\boldsymbol{\Delta}}[k]}{\|\widetilde{\boldsymbol{\Delta}}[k]\|}\sin\left(\|\widetilde{\boldsymbol{\Delta}}[k]\|t_{\mathrm{opt}}\right) \quad (23)$$

Results obtained in (17), (21), (22) and (23) are consistent with those reported in [18]. That is, if we let $M = 1$ in our proposed GSPC scheme, the resulted outputs exactly match with those given in [18], while on the contrary, it is not straightforward (i.e., let $M > 1$ in [18]). Therefore, we claim that the prediction framework studied in [18] is actually a special case of our proposed GSPC for rank-1 transmission. In the following, some mathematical interpretations on the proposed quantization scheme (i.e., (13)) are provided assuming transmit beamforming.[5] We first denote $\tilde{\rho}_i = \boldsymbol{W}^H[k-1]\boldsymbol{W}_i$ and $\tilde{d}_i = \sqrt{1-|\tilde{\rho}_i|^2}$. Then we rewrite $\widetilde{\boldsymbol{\Delta}}_i$ and (12) accordingly as

$$\widetilde{\boldsymbol{\Delta}}_i = \tan^{-1}\left(\frac{\tilde{d}_i}{|\tilde{\rho}_i|}\right)\frac{\tilde{\rho}_i^*\boldsymbol{W}_i - \boldsymbol{W}[k-1]}{\tilde{d}_i}, \quad (24)$$

$$\widetilde{\boldsymbol{W}}_i = \boldsymbol{W}_i\cos(\|\widetilde{\boldsymbol{\Delta}}_i\|t_i) + \frac{\widetilde{\boldsymbol{\Delta}}_i}{\|\widetilde{\boldsymbol{\Delta}}_i\|}\sin(\|\widetilde{\boldsymbol{\Delta}}_i\|t_i). \quad (25)$$

[5]We note that our analysis is suited for spatial multiplexing as well. Using the transmit beamforming as the example to illustrate is for simplicity reason.

By substituting (23), (24) and (25) into (13), and after some simple manipulations, we obtain (26)-(29). For slow-varying channels, we have $\epsilon \approx 1$, $\mu \approx 0$ and $\eta \approx 0$. This results in the following approximation, given as

$$\max_{\boldsymbol{W}_i\in\mathcal{W}}\left|\widehat{\boldsymbol{U}}^H[k+1]\widetilde{\boldsymbol{W}}_i\right|^2 \triangleq \max_{\boldsymbol{W}_i\in\mathcal{W}}\left|\boldsymbol{U}^H[k]\boldsymbol{W}_i\right|^2. \quad (30)$$

(30) indicates that under this scenario, the proposed quantization criterion is approximately the same as for the block fading channel model. Relatively, for fast-varying channels, (27) implies that the selected codeword $\boldsymbol{W}[k]$ should not only be as "close" as to the estimated CSI $\boldsymbol{U}[k]$ (the first term of the RHS of (27)), the direction of its transported tangent vector should also point to the optimally predicted CSI $\widehat{\boldsymbol{U}}[k+1]$ as accurate as possible (the second term of the RHS of (27)), with some modifying and scaling factors (ϵ, μ, η).

V. Simulation Results

In this section, we simulate the memoryless Grassmannian subspace packing (GSP), the GST algorithm and the linear minimum mean squared error (MMSE) channel predictor for comparison with our proposed method. $N_b = \log_2 L$ denotes the number of feedback bits. In addition, regarding our proposed GSPC algorithm, the optimized step size parameter is quantized by using a 10-bit uniform scalar codebook, and fed back to the transmitter every ten channel uses. Time-varying channels with memories are obtained according to the AR1 fading model. The correlation coefficient of the AR1 fading model is defined as $\alpha = J_0(2\pi\beta)$, where J_0 is the Bessel function of zeroth order while β represents the normalized Doppler frequency. At time $k+1$, the correlated MIMO channel is expressed as $\boldsymbol{H}[k+1] = \alpha\boldsymbol{H}[k] + \sqrt{1-\alpha^2}\boldsymbol{G}[k]$ where $\boldsymbol{G}[k]$ represents a $M_r \times M_t$ complex matrix with each entry distributed according to $\mathcal{CN}(0,1)$. Perfect channel estimation is assumed at the receiver.

The *received power* $\left\|\boldsymbol{H}^H[k+1]\widetilde{\boldsymbol{W}}[k+1]\right\|_F^2$ can be considered as an effective metric to measure the robustness of the proposed precoding scheme [17]. In Fig. 1, the cumulative density function (CDF) of the *received power* is plotted. From Fig. 1, we observe that the proposed GSPC encoding/decoding algorithm significantly outperforms the existing prediction techniques and approaches that of the optimal prediction. This gain increases with increase in the system configurations. For instance, at $M_t = M_r = 4, M = 2$, the performance gap between our proposed approach and the linear MMSE channel predictor is 1dB at the CDF level of 10^{-3}. While this gain grows to nearly 2dB with the configurations of $M_t = 8, M_r = 4, M = 3$.

The CDF can be interpreted as the first-order statistics of the *received power*. In this example, the level crossing rate (LCR) of the *received power* is evaluated, which can be considered as the second-order statistics. The LCR reveals the expected rate at which the signal envelope crosses a given level in the downward direction. Simulation results shown in Fig. 2 further validate the effectiveness and the robustness of the proposed

$$\max_{\boldsymbol{W}_i \in \mathcal{W}, i=1,\cdots,L} \left| \widehat{\boldsymbol{U}}^H[k+1] \widetilde{\boldsymbol{W}}_i \right|^2 \triangleq \max_{\boldsymbol{W}_i \in \mathcal{W}, i=1,\cdots,L} \left| \epsilon \boldsymbol{U}^H[k] \boldsymbol{W}_i + \mu \frac{\widetilde{\boldsymbol{\Delta}}^H[k]}{\|\widetilde{\boldsymbol{\Delta}}[k]\|} \frac{\widetilde{\boldsymbol{\Delta}}_i}{\|\widetilde{\boldsymbol{\Delta}}_i\|} + \eta \right|^2, \tag{26}$$

where

$$\epsilon = \left[\cos(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i) + \frac{\tilde{\rho}_i^*}{\tilde{d}_i} \sin(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i) \right] \cos(\|\widetilde{\boldsymbol{\Delta}}[k]\| t_{\text{opt}}), \tag{27}$$

$$\mu = \left[\sin(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i) + \frac{\tilde{d}_i}{\tilde{\rho}_i^*} \cos(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i) \right] \sin(\|\widetilde{\boldsymbol{\Delta}}[k]\| t_{\text{opt}}), \tag{28}$$

$$\eta = \frac{1}{\tilde{\rho}_i^*} \frac{\widetilde{\boldsymbol{\Delta}}^H[k]}{\|\widetilde{\boldsymbol{\Delta}}[k]\|} \boldsymbol{W}[k-1] \sin(\|\widetilde{\boldsymbol{\Delta}}[k]\| t_{\text{opt}}) \cos(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i) - \frac{1}{\tilde{d}_i} \boldsymbol{U}^H[k] \boldsymbol{W}[k-1] \cos(\|\widetilde{\boldsymbol{\Delta}}[k]\| t_{\text{opt}}) \sin(\|\widetilde{\boldsymbol{\Delta}}_i\| t_i). \tag{29}$$

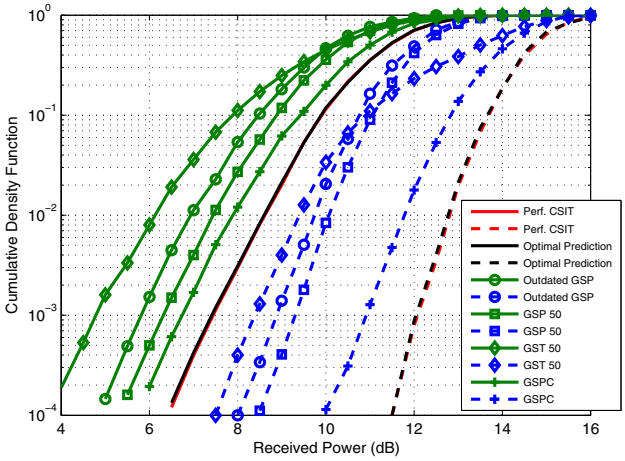

Fig. 1. Cumulative density function (CDF) of the received power; solid lines – $\{M_t = M_r = 4, M = 2\}$; dashed lines – $\{M_t = 8, M_r = 4, M = 3\}$; $N_b = 5$; $\beta = 0.01$; GSP 50 denotes a length-50 linear MMSE channel predictor using GSP codebooks for quantization; GST 50 implies that a length-50 linear MMSE channel predictor is employed as part of the entire tracking algorithm.

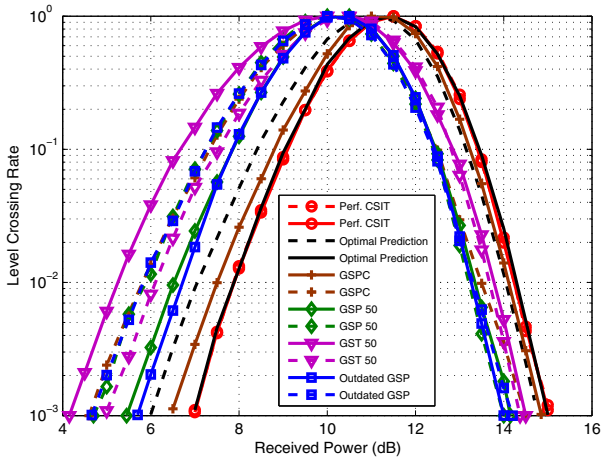

Fig. 2. Level crossing rate (LCR) of the received power; $M_t = M_r = 4$; $M = 2$; $N_b = 5$; solid lines – $\{\beta = 0.01\}$; dashed lines – $\{\beta = 0.1\}$.

GSPC algorithm, and are consistent with those presented in Fig. 1.

In addition to the *received power*, the chordal distance between the observed CSI $\widetilde{\boldsymbol{U}}[k+1]$ and the representative quantized CSI $\widetilde{\boldsymbol{W}}[k+1]$ can be used to characterize the accuracy of the prediction with quantization errors. Here, the chordal distance between two subspaces is defined as [21]

$$d_c = \frac{1}{2} \left\| \widetilde{\boldsymbol{U}}[k+1] \widetilde{\boldsymbol{U}}^H[k+1] - \widetilde{\boldsymbol{W}}[k+1] \widetilde{\boldsymbol{W}}^H[k+1] \right\|_F. \tag{31}$$

In Fig. 3, the MSE of the chordal distance is plotted, which is calculated as[6]

$$E\left[\frac{1}{2} \left\| \widetilde{\boldsymbol{U}}[k+1] \widetilde{\boldsymbol{U}}^H[k+1] - \widetilde{\boldsymbol{W}}[k+1] \widetilde{\boldsymbol{W}}^H[k+1] \right\|_F^2 \right]. \tag{32}$$

[6](31) and (32) are expressed by using the proposed GSPC algorithm. They can be similarly defined for other transmission schemes.

From the evaluation results, we can see that in the region of low-to-medium normalized Doppler frequencies (i.e., $\beta \leq 0.1$), the proposed GSPC algorithm exhibits the best MSE performance. When β exceeds 0.1, the GST algorithm shows superior performance over the proposed approach in terms of the MSE performance. This phenomena is observed for both $N_b = 5$ and $N_b = 7$. However, it is worth noting here that $\beta > 0.1$ rarely occurs in typical cellular systems.

In this example, transmit beamforming is evaluated in terms of the uncoded BER performance assuming QPSK modulation scheme. In addition to our proposed algorithm, the memoryless quantization approach, the algorithm presented in [18] and the linear MMSE channel predictor are also simulated for comparison. Again, our proposed approach exhibits the best error rate performance and shows nearly the same diversity order as compared to the case that employs the optimally predicted beamforming vector. Especially at $\beta = 0.2$, the gap between the proposed scheme and the optimal prediction nearly vanishes. Furthermore, the algorithm proposed in [18] significantly outperforms the memoryless quantization scheme

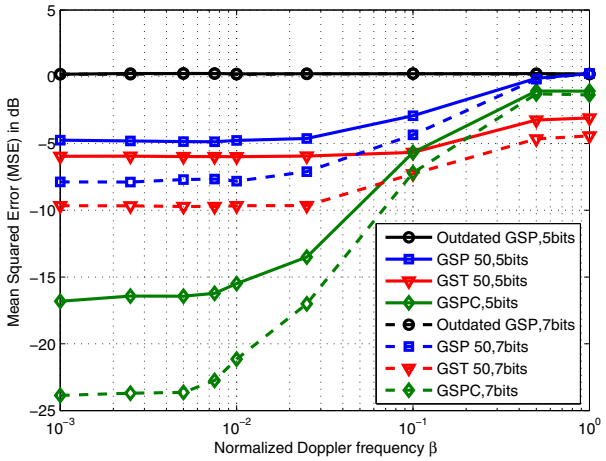

Fig. 3. Mean squared error (MSE) of the chordal distance versus the normalized Doppler frequency β; $M_t = M_r = 4$; $M = 2$.

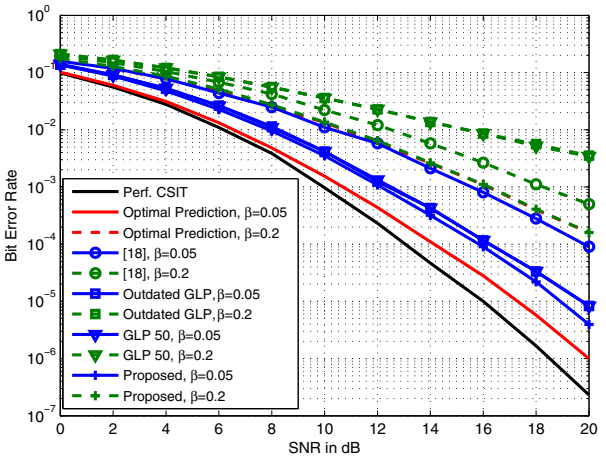

Fig. 4. Bit error rate of uncoded QPSK; GLP is short for Grassmannian line packing; GLP 50 denotes a length-50 linear MMSE channel predictor using GLP codebooks for quantization; $M_t = 4$; $M_r = 1$; $M = 1$; $N_b = 5$.

at $\beta = 0.2$, but collapses at $\beta = 0.05$. This implies that at relatively low feedback rate, the algorithm proposed in [18] is not robust to the variations in the mobility environment.

VI. Conclusions

In this paper, by exploiting the geometric properties of the Grassmannian manifold, we have invoked a novel prediction framework to compensate for the feedback delay in closed-loop precoded MIMO systems. The proposed encoding/decoding algorithm is propagation model independent and has no constraint on the transmission rank. We conclude through simulations that our proposed GSPC scheme exhibits superior system performances with respect to the existing prediction techniques, for both spatial multiplexing and transmit beamforming MIMO systems, using the same amount of feedback bits.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," in *Wireless Personal Communications*, vol. 6, Mar. 1998, pp. 311–335.

[2] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," in *IEEE Journal on Selec. Areas in Comm.*, vol. SAC-5, 1987, pp. 871–878.

[3] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," in *IEEE Journal on Selec. Areas in Comm.*, vol. 26, 2008, pp. 1341–1365.

[4] B. Mondal and R. W. Heath, "Quantization on the Grassmannian manifold," in *IEEE Trans. on Signal Process.*, vol. 55, no. 8, Aug. 2007, pp. 4208–4216.

[5] D. J. Love, R. W. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," in *IEEE Trans. on Info. Theo.*, vol. 49, Oct. 2003, pp. 2735–2747.

[6] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," in *IEEE Trans. on Info. Theo.*, vol. 51, Aug. 2005, pp. 2967–2976.

[7] T. Inoue and R. W. Heath, "Kerdock codes for limited feedback precoded MIMO systems," in *IEEE Trans. on Signal Process.*, vol. 57, no. 9, Sep. 2009, pp. 3711–3716.

[8] K. Huang, R. W. Heath, and J. G. Andrews, "Limited feedback beamforming over temporally-correlated channels," in *IEEE Trans. on Signal Process.*, vol. 57, May 2009, pp. 1959–1975.

[9] B. C. Banister and J. R. Zeidler, "Feedback assisted transmission subspace tracking for MIMO systems," in *IEEE Journal on Selec. Areas in Comm.*, vol. 21, Apr. 2003, pp. 452–463.

[10] J. C. Roh and B. D. Rao, "Efficient feedback methods for MIMO channels based on parameterization," in *IEEE Trans. on Wireless Comm.*, vol. 6, Jan. 2007, pp. 282–292.

[11] C. Simon and G. Leus, "Feedback reduction for spatial multiplexing with linear precoding," in *Proc. IEEE Int. Conf. Acoust., Speech and Signal Proc.*, vol. 3, Apr. 2007.

[12] S. Zhou and G. B. Giannakis, "How accurate channel prediction needs to be for transmit-beamforming with adaptive modulation over Rayleigh MIMO channels," in *IEEE Trans. on Wireless Comm.*, vol. 3, Jul. 2004, pp. 1285–1294.

[13] N. Benvenuto, E. Conte, S. Tomasin, and M. Trivellato, "Predictive channel quantization and beamformer design for MIMO-BC with limited feedback," in *IEEE Trans. on Signal Process.*, vol. 57, May 2009, pp. 1959–1975.

[14] T. R. Ramya and S. Bhashyam, "Eigen-beamforming with delayed feedback and channel prediction," in *IEEE ISIT*, Jun. 2009.

[15] L. Liu and H. Jafarkhani, "Novel transmit beamforming schemes for time-selective fading multi-antenna systems," in *IEEE Trans. on Signal Processing*, vol. 54, Dec. 2006, pp. 4767–4781.

[16] ——, "Successive transmit beamforming algorithms for multiple-antenna OFDM systems," in *IEEE Trans. on Wireless Comm.*, vol. 6, Apr. 2007, pp. 1512–1522.

[17] J. Yang and D. B. Williams, "Transmission subspace tracking for MIMO systems with low-rate feedback," in *IEEE Trans. on Comm.*, vol. 55, no. 8, Aug. 2007, pp. 1629–1639.

[18] T. Inoue and R. W. Heath, "Predictive coding on the Grassmannian manifold," in *submitted to IEEE Trans. on Signal Process.*, Aug. 2009.

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

[20] J. C. Roh and B. D. Rao, "Transmit beamforming in multiple-antenna systems with finite rate feedback: a VQ-based approach," in *IEEE Trans. on Info. Theo.*, vol. 52, no. 3, Mar. 2006, pp. 1101–1112.

[21] D. J. Love and R. W. Heath, "Limited feedback unitary precoding for orthogonal space-time block codes," in *IEEE Trans. Signal Processing*, vol. 53, no. 1, Jan. 2005, pp. 64–73.