Incremental Grassmannian Feedback Schemes for Multi-User MIMO Systems

Ahmed Medra and Timothy N. Davidson

Abstract—The communication of side information forms a key component of several effective strategies for transmitter adaptation to slowly fading channels. When the relevant side information is a subspace, the feedback scheme can be viewed as a lossy source compression scheme on the Grassmannian manifold. Memoryless vector quantization on each fading block is a viable compression scheme, but it neglects any temporal correlation between the blocks. In this paper, we propose an incremental approach to Grassmannian quantization that takes advantage of temporal correlation. The approach leverages existing codebooks for memoryless quantization schemes and employs a quantized form of geodesic interpolation. Two schemes that implement the principles of the proposed approach are presented. In the first scheme, the choice of the step size in the incremental update is adapted to a first-order GaussMarkov model for the channel, which enables the use of higher resolution codebooks. In the second scheme, a single bit is allocated to the step size, which enables adaptation of the step size to the channel realization rather than the channel statistics. This provides substantial robustness against mismatches in the model for the temporal correlation. A distinguishing feature of the proposed approach is that the direction of the geodesic interpolation is specified implicitly using a point in a conventional codebook. As a result, the approach has an inherent ability to recover autonomously from errors in the feedback path. Simulation results demonstrate that these features result in improved performance over some existing schemes in a variety of channel environments.

Index Terms—Channel temporal correlation, downlink, geodesic, limited feedback.

I. Introduction

IN the design of systems for communicating over slowly fading channels, the provision of information of the state of the channel to the transmitter offers the opportunity to enhance a variety of measures of the quality of service provided to the receiver(s) by enabling the adaptation of the transmission to the channel state [1]. In particular, in downlink systems in which the transmitter, and possibly the receivers, have multiple antennas, the availability of the channel state information (CSI) at the transmitter enables effective management of the interference incurred through simultaneous transmissions to multiple users in the same band [2]. As a result, schemes for providing the transmitter with various forms of CSI form core components of standards such as WiMax [3] and LTE-Advanced [4], [5] and envisioned schemes for transmitter coordination [6] and Massive MIMO [7]. In time division duplex (TDD) systems, the transmitter may be able to obtain sufficiently accurate CSI on the reverse link, but in frequency division duplex (FDD) systems, the CSI is estimated at the receivers and then fed back to the transmitter. In such systems, the receivers typically perform a form of lossy source compression [8] on the information to be fed back to the transmitter in order to manage the allocation of communication resources to the feedback task [9]. (TDD systems may also benefit from the provision of feedback [10].) Once the transmitter has obtained its estimates of the channel(s), it determines a transmission strategy and informs the receivers. In the scenarios that we will consider that process will include a "dedicated" training phase [11] that is akin to that in LTE-Advanced [5].

In many CSI feedback schemes for systems with multiple antennas, the development of an effective lossy source compression scheme can be simplified by partitioning the information to be quantized. For example, for multiple-input multiple-output (MIMO) point-to-point links, the receiver may inform the transmitter separately of the directions in which it should transmit and the power that it should allocate to each direction. Similarly, for the MIMO downlink, each receiver may inform the transmitter of the directions of its channels and a measure of the "quality" of the channel in each direction. In some applications, further simplification is possible. In particular, in single user MIMO systems that employ uniform power loading, many communication objectives depend only on the subspaces spanned by the directions of transmission, rather than the directions themselves; e.g., [12], [13]. Similarly, in a number of simplified multi-user MIMO systems, the design of the transmitter is dependent on the subspace spanned by the channel [14]. Since subspaces of a given dimension can be represented by points on the corresponding Grassmannian manifold [15], this manifold is the natural setting for the source compression problem. In this paper, we will focus on Grassmannian feedback schemes, but the principles of our approach can be applied to systems that feed back information regarding specific directions, rather

than subspaces, by performing lossy source compression on the Stiefel manifold.

For the uncorrelated i.i.d. block fading Rayleigh channel, memoryless uniform vector quantization on the Grassmannian manifold is an effective strategy that can offer substantial performance gain at the price of only few feedback bits [12], [13]. In this setting, a quantization codebook is designed offline and known to all terminals [16]–[21]. For each fading block, the receiver sends to the transmitter the index of one element of the codebook, which represents the quantized version of the subspace. Since it is memoryless, this scheme neglects any correlation that may exist between fading blocks in practice, and hence it may require a larger amount of feedback than necessary. However, neglecting temporal correlation provides inherent robustness against mismatches in any model for temporal correlation.

Several methods have been proposed for developing limited feedback schemes that take advantage of temporal correlation between fading blocks; e.g., [22]–[32]. A number of these schemes are based on a differential quantization strategy in which the scheme seeks to incrementally update the current estimate of the subspace. Typically, these schemes involve two codebooks, the first of which is a conventional codebook designed for a memoryless quantization scheme and is used to initialize the scheme. The second codebook (or family of codebooks) is used to quantize local changes around the current subspace. A variety of approaches have been proposed for designing the second codebook. One approach is to construct a simple parameterized codebook for the neighborhood of a point on the manifold [23]–[26]. Since the geometry of the Grassmannian manifold is analogous to that of a sphere, these local codebooks are often called spherical or polar cap codebooks. In a number of schemes of this kind, the codebook consists of a set of points that are carefully distributed on a ring. In some schemes, the radius of the ring is scaled according to the channel statistics, while in others, the radius is successively shrunk. In some schemes, the spherical cap codebook is rotated each time it is used. An alternative to the spherical/polar cap approaches, is to employ a codebook of tangents to the Grassmannian manifold. Once the transmitter receives the index of the tangent, it can construct the updated version of the subspace by taking a step of a specified size on the manifold in a direction specified by the tangent [27], [28]. The geometric constructions in these differential methods can be extended to develop predictive quantization schemes [29], [30]. Finally, guidelines for selecting the feedback update period and the number of feedback bits in differential quantization schemes for beamforming based systems were derived in [31].

In this paper, we propose an incremental feedback approach to the Grassmannian quantization problem that exploits the temporal correlation between fading blocks and yet requires only a single codebook that is designed using conventional techniques for memoryless Grassmannian quantization, such as those in [16]–[21]. Similar to the principles of the methods in [22], [27], [29], we propose to update the current estimate of the subspace by taking a step along a specified geodesic on the manifold. The proposed approach differs in that the direction is captured using a conventional Grassmannian codebook and the step size

along the geodesic can be quantized in a variety of ways. In order to implement the principles of the proposed approach, we develop two feedback schemes that differ in the way in which the step size is communicated. The first scheme is a specialization of the incremental approach that is tailored to a first-order Gauss-Markov model with a known correlation coefficient. In this *model-based incremental scheme*, the step size is recursively updated in a manner that is dependent on the correlation coefficient. As a result, no bits need to be assigned to feed back the step size and hence a conventional Grassmannian codebook with higher resolution can be used. However, the performance of the model-based scheme is dependent on the accuracy with which the correlation coefficient is estimated and tracked.

In the second scheme, conventional Grassmannian codebooks of lower resolution are used and the remainder of the bit budget is used to adapt the step size to the channel realization (rather than to the channel statistics). Even with only one bit allocated to the step size, this scheme provides substantial robustness to mismatches in the presumed statistical model and hence we will refer to this scheme as the *robust incremental scheme*. An interesting feature of both of the proposed incremental schemes is that they exhibit an intrinsic property that allows them to recover from errors caused by the feedback channel. In the case of the robust scheme, this recovery property can be considered to be a "self reseting" feature, in which the transmitter and the receiver can perfectly re-synchronize to each other after a feedback error.

The performance of the proposed schemes will be compared to that of other existing schemes in the cases of single user (SU) and single-cell multi-user (MU) MIMO systems. It is worth mentioning that the proposed schemes can be also implemented in more complicated scenarios, such as the multi-cell downlink, but the underlying principles for channel quantization in that setting are similar to those for the case of MU-MIMO systems. In the case of SU-MIMO systems, having CSI at the transmitter side can improve the achievable rate, but it is not necessary to achieve the multiplexing gain of the system. The case of MU-MIMO is quite different, in that the multiplexing gain provided by the channel cannot be achieved in the absence of CSI at the transmitter [33]. In the simulation section, we will evaluate the performance of the proposed incremental schemes in several channel scenarios and system models. For example, we compare the achievable rate in the case of SU-MIMO with perfect channel estimation and with channel model mismatches. The results show the improved performance of the proposed schemes in comparison with some existing ones. In the case of MU-MIMO, we adopt the Zero-Forcing Beamforming and Block-diagonalization systems and the results show the performance gains in the sum rate achieved by our schemes, especially with estimation errors in the channel model. Finally, we evaluate the performance of the proposed scheme in the case of feedback error and we show that our schemes have an inherent ability to recover from feedback errors and provide performance gains, in contrast to other existing schemes.

II. SYSTEM MODEL

In order to focus on the principles of the proposed schemes, we will present our results in case of SU and MU-MIMO sys-

tems. We will consider a MU-MIMO downlink system with a base station with M_t transmit antennas communicating to K users, the kth of which has M_{rk} receive antennas. The received signal for the kth user is given by:

$$\mathbf{y}_k = \sqrt{\frac{E_s}{Q}} \mathbf{H}_k \mathbf{T} \mathbf{s} + \mathbf{z}_k \qquad (1)$$

where $\mathbf{H}_k \in \mathbb{C}^{M_{rk} \times M_t}$ is the channel matrix from the transmitter to the kth user, $Q \leq \min(M_t, \sum_k M_{rk})$ is the number of transmitted data streams, and $\mathbf{T} \in \mathbb{C}^{M_t \times Q}$ is the transmitter preprocessing matrix which is normalized so that $\mathrm{trace}(\mathbf{T}^H \mathbf{T}) = Q$. The transmitted symbols form the vector $\mathbf{s}$, which is multiplied by the preprocessing matrix $\mathbf{T}$ producing a signal vector $\mathbf{x} = \sqrt{\frac{E_s}{Q}} \mathbf{T} \mathbf{s}$, where E_s is the total transmit energy assuming that $\mathrm{E}\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}_Q$. Finally $\mathbf{z}_k$ is the additive noise at user k which is assumed to be zero-mean complex Gaussian with normalized covariance, $\mathrm{E}\{\mathbf{z}_k \mathbf{z}_k^H\} = \mathbf{I}_{M_{r_k}}$.

We consider a scenario in which the channel changes in a block fading manner with temporal correlation between blocks and we assume that each receiver has perfect knowledge of its channel matrix. The receiver seeks to provide information about this channel to the base station, and we consider the case in which the receivers employ partitioned quantization schemes in which one of the partitions involves the quantization of a subspace. Although the focus of this paper is on developing quantization schemes with memory that can exploit the temporal correlation between blocks, we first briefly review some examples of memoryless quantization schemes for such applications.

A. Single-User MIMO (SU-MIMO)

In the case of a point-to-point MIMO link, for which $K = 1$, a simplified limited feedback strategy that has proven effective in practice is for the receiver to inform the transmitter of the directions in which transmission should take place and for the transmitter to allocate power uniformly over these directions. In that case, many communication objectives depend on the subspace spanned by the directions and not on the directions themselves, and hence quantization on the Grassmannian manifold arises naturally. If the transmitter and the receiver are provided with a codebook $\mathcal{F} = \{\mathbf{F}_i\}_{i=1}^L$ containing $L = 2^B$ tall matrices of size $M_t \times Q$ with orthonormal columns, each of which is the Grassmannian representative of the subspace that it spans, then the receiver selects the matrix (codeword) $\mathbf{F}_{i^\star}$ that maximizes a chosen performance metric and sends the index $i^\star$ to the transmitter so it can identify this subspace [13]. For example, if the performance metric is the Gaussian mutual information, then the receiver selects a codeword $\mathbf{F}_i \in \mathcal{F}$ according to

$$\mathbf{F}_{i^\star} = \arg \max_{\mathbf{F}_i \in \mathcal{F}} \log \det \left(\mathbf{I}_{M_r} + \frac{E_s}{Q} \mathbf{F}_i^H \mathbf{H}^H \mathbf{H} \mathbf{F}_i \right). \qquad (2)$$

The index of $\mathbf{F}_{i^\star}$ is send back to the transmitter which will use $\mathbf{F}_{i^\star}$ as the current preprocessing matrix, i.e., $\mathbf{T} = \mathbf{F}_{i^\star}$. Therefore, the achievable rate is the optimal value of the problem in (2). That rate, or the index of an appropriate coding and modulation scheme, can be also fed back to the transmitter.

B. Single Receive Antenna Downlink (MU-MISO)

In a multi-user downlink scenario in which each user has a single receive antenna ($M_{r_k} = 1$ and hence $Q = K$), the feedback of the channel state information is often partitioned into a channel direction vector and a scalar that represents the "quality" of the channel. Similar to the single user case, the transmitter and the receivers are provided with a codebook $\mathcal{F} = \{\mathbf{f}_i\}_{i=1}^L$ of unit norm vectors and the kth receiver selects the codeword (vector) that solves [34]

$$\mathbf{f}_{i^\star,k} = \arg \max_{\mathbf{f}_i \in \mathcal{F}} |\mathbf{h}_k \mathbf{f}_i^*|, \qquad (3)$$

and transmits $i^\star$ back to the transmitter. The transmitter uses these indices to reconstruct $\mathbf{f}_{i^\star,k}$ for all k, and then uses these vectors to calculate the (normalized) preprocessing matrix $\mathbf{T}$ in (1). A subsequent "dedicated training" step (e.g., [11]) enables the kth receiver to determine $\mathbf{h}_k \mathbf{T}$, and the achievable rate for the kth user, which treats interference as noise, is given by

$$R_k^{BF} = \log \left(1 + \frac{\frac{E_s}{K} |\mathbf{h}_k \mathbf{t}_k|^2}{1 + \sum_{j \neq k} \frac{E_s}{K} |\mathbf{h}_k \mathbf{t}_j|^2} \right) \qquad (4)$$

In the case that the transmitter performs Zero-Forcing Beamforming (ZFBF)[14] with equal power loading, a matrix $\tilde{\mathbf{T}}$ is constructed by aggregating the quantized channel directions sent by the scheduled users according to:

$$\tilde{\mathbf{T}} = [\mathbf{f}_{i^\star,1}, \mathbf{f}_{i^\star,2}, \ldots, \mathbf{f}_{i^\star,K}]^\dagger, \qquad (5)$$

where $\mathbf{A}^\dagger$ is the pseudo inverse of $\mathbf{A}$. The precoding matrix $\mathbf{T}^{ZFBF}$ is then formed by normalizing each column in $\tilde{\mathbf{T}}$, and if $\mathbf{t}_j$ denotes the jth column of $\mathbf{T}$, then $\mathbf{f}_{i^\star,k}^H \mathbf{t}_j^{ZFBF} = 0$ for all $k \neq j$. We observe that if two users select the same $\mathbf{f}_{i^\star} \in \mathcal{F}$, then the matrix $\mathbf{T}$ will have rank less than K. In practice, this scenario is avoided by providing each user a uniquely rotated version of the codebook $\mathcal{F}$. For comparison, in the case of perfect channel state information at the transmitter ZFBF results in $\mathbf{t}_j^{ZFBF}$ being orthogonal to $\mathbf{h}_k$ for $j \neq k$ and hence in that case the achievable rate simplifies to $R_k^{ZFBF-CSIT} = \log \left(1 + \frac{E_s}{K} |\mathbf{h}_k \mathbf{t}_k^{ZFBF}|^2 \right)$.

C. Multiple Receive Antenna Downlink (MU-MIMO)

In this scenario, each receiver has more than one antenna and hence can receive more than one data stream. Similar to the single antenna downlink case, the feedback strategy involves sending information about the channel direction information and the quality of the channel. Assuming that a Grassmannian codebook $\mathcal{F} = \{\mathbf{F}_i\}_{i=1}^L$ is known to the transmitter and the receiver, the kth receiver selects the index of a codeword according to [35]

$$\mathbf{F}_{i^\star,k} = \arg \max_{\mathbf{F}_i \in \mathcal{F}} \|\mathbf{H}_k \mathbf{F}_i\|_F^2, \qquad (6)$$

where $\| \cdot \|_F$ denotes the Frobenius norm, and sends $i^\star$ to the transmitter. The transmitter reconstructs $\mathbf{F}_{i^\star,k}$, for all k, from the received indices and then constructs the precoding matrix $\mathbf{T} = [\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_K]$, where $\mathbf{T}_k \in \mathbb{C}^{M_t \times Q_k}$ and $Q_k \leq \min\{M_t, M_{r_k}\}$ is the number of streams transmitted to the kth

receiver. Following the dedicated training step the kth receiver can determine $\mathbf{H}_k\mathbf{T}$, and the achievable rate for the kth user, which treats interference as noise, is given by

$$R_k = \log\left(\frac{\det\left(\mathbf{I}_{M_{r_k}} + \sum_j \frac{E_s}{Q_j}\mathbf{T}_j^H\mathbf{H}_k^H\mathbf{H}_k\mathbf{T}_j\right)}{\det\left(\mathbf{I}_{M_{r_k}} + \sum_{j\neq k}\frac{E_s}{Q_j}\mathbf{T}_j^H\mathbf{H}_k^H\mathbf{H}_k\mathbf{T}_j\right)}\right) \quad (7)$$

In the case that the transmitter performs Block Diagonalization (BD) with equal power loading [35], each $\mathbf{T}_k$ is chosen such that $\mathbf{F}_{i^\star,j}\mathbf{T}_k^{BD} = 0$, for all $j \neq k$, and so that its columns have unit norm. This can be formed by finding an orthonormal basis for the null space of the matrix constructed by stacking all the $\mathbf{F}_{i^\star,j}$ matrices, $j \neq k$, together. For comparison, in the case of perfect channel state information at the transmitter, $\mathbf{H}_j\mathbf{T}_k^{BD} = 0$, for all $j \neq k$, and hence the achievable rate simplifies to $R_k^{BD-CSIT} = \log\det\left(\mathbf{I}_{M_{r_k}} + \frac{E_s}{Q_k}\mathbf{T}_k^{BD^H}\mathbf{H}_k^H\mathbf{H}_k\mathbf{T}_k^{BD}\right)$.

III. TOPOLOGY OF THE GRASSMANNIAN MANIFOLD

Memoryless quantization schemes, such as those described above, are effective schemes for block fading channel models in which the blocks are independent. In scenarios in which the blocks are correlated, memoryless quantization schemes can still be employed, but schemes that seek to exploit that correlation have the potential to reduce the quantization error (and implicitly its impact on the system performance) or reduce the amount of feedback that is required to achieve a given level of performance. The approach proposed in this paper for exploiting correlation between blocks is based on the topology of the Grassmannian manifold [15], [36], [37], which we now briefly review.

The Grassmannian manifold $\mathbb{G}_{M_t,M}$ is a representation of subspaces of dimension M in $\mathbb{C}^{M_t}$. Such a subspace can be described as the linear span of an orthonormal basis, and that basis can be captured by the columns of a matrix $\mathbf{X} \in \mathbb{C}^{M_t \times M}$ that satisfies $\mathbf{X}^H\mathbf{X} = \mathbf{I}_M$. As there is a continuum of matrices that can represent a given subspace, these matrices can be deemed to form an equivalence class in terms of representing subspaces. Each "point" on the Grassmannian manifold $\mathbb{G}_{M_t,M}$ is a single matrix $\mathbf{X}$ that represents this equivalence class for the given subspace.

The intuition behind the proposed incremental schemes depends on the concept of the geodesic (the shortest path) between two points on the manifold. To write an equation for the geodesic from the point $\mathbf{F}_i$ on the manifold to the point $\mathbf{F}_j$, we will use the notion of the tangent to the manifold at the point $\mathbf{F}_i$ in the direction of the point $\mathbf{F}_j$, which can be written as

$$\Delta(\mathbf{F}_i, \mathbf{F}_j) = (\mathbf{I} - \mathbf{F}_i\mathbf{F}_i^H)\mathbf{F}_j\left(\mathbf{F}_i^H\mathbf{F}_j\right)^{-1}. \quad (8)$$

If we let $\Delta(\mathbf{F}_i, \mathbf{F}_j) = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$ denote the compact singular value decomposition of $\Delta(\mathbf{F}_i, \mathbf{F}_j)$ and define $\mathbf{\Phi} = \tan^{-1}(\mathbf{\Sigma})$, with $\tan^{-1}(\cdot)$ being defined elementwise, then the points on the geodesic from $\mathbf{F}_i$ to $\mathbf{F}_j$ can be written as

$$\mathbf{F}(t) = \mathbf{F}_i\mathbf{V}\cos(\mathbf{\Phi}t)\mathbf{V}^H + \mathbf{U}\sin(\mathbf{\Phi}t)\mathbf{V}^H \quad (9)$$

where $t \in [0,1]$ tracks the progress along the geodesic from $\mathbf{F}_i$ to $\mathbf{F}_j$. In the case of beamforming, where the
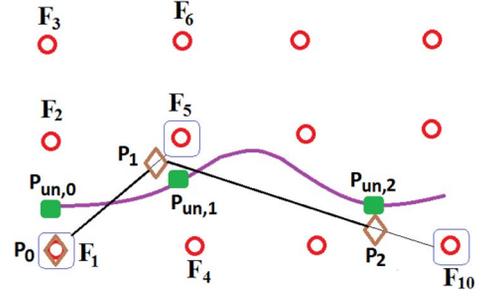
Fig. 1. A pictorial representation of the proposed technique. The points marked $\mathbf{P}_{\mathrm{un},n}$ denote the sequence of points on the manifold that would be chosen if the feedback were unlimited. The circles denote the points $\mathbf{F}_i$ in the memoryless Grassmannian codebook (which do not have to be uniformly spaced), and the points $\mathbf{P}_n$ denote the points generated by the proposed technique. Note that the Grassmannian manifold is compact, and hence the edge effects that appear in this pictorial representation do not arise in practice.

points on the manifold are elements of $\mathbb{C}^{M_t}$, (9) can be simplified to $\mathbf{f}(t) = -\mathbf{f}_i\cos(\phi t) + \frac{\Delta(\mathbf{f}_i,\mathbf{f}_j)}{\|\Delta(\mathbf{f}_i,\mathbf{f}_j)\|}\sin(\phi t)$, where $\phi = \tan^{-1}(\|\Delta(\mathbf{f}_i,\mathbf{f}_j)\|)$.

IV. PRINCIPLES OF THE PROPOSED INCREMENTAL FEEDBACK SCHEMES

The proposed approach to incremental Grassmannian feedback is based on the observation from (9) that the current representative of the subspace can be updated by taking a step of a specified size along the geodesic in the direction of a point selected from a conventional Grassmannian codebook. To describe that approach, we will assume that B bits that are available for feedback, and that those bits are partitioned into two sets of size B_{cb} and B_{step}, respectively. The set of B_{cb} bits is used to index the elements of a Grassmannian codebook $\mathcal{F}_{\mathrm{cb}}$ of size $L_{\mathrm{cb}} = 2^{B_{\mathrm{cb}}}$ from a memoryless scheme, and the set of B_{step} bits is used to index a quantization of the interval $[0,1]$, $\mathcal{T} = \{\bar{t}_1, \ldots, \bar{t}_{L_{\mathrm{step}}}\}$, where $L_{\mathrm{step}} = 2^{B_{\mathrm{step}}}$. The basic operation of the proposed scheme is illustrated in Fig. 1. In that figure, the points $\{\mathbf{P}_{un,n}\}$ denote the sequence of points on the manifold that would be chosen in case of unlimited feedback. That is, the true channel direction information (CDI) of the channel.[1] The point $\mathbf{P}_n$ denotes the quantized version of the CDI for block n under the proposed limited feedback model and the points $\{\mathbf{F}_i\}$ denote the elements of the conventional Grassmannian codebook $\mathcal{F}_{\mathrm{cb}}$. The initial quantized CDI $\mathbf{P}_0$ is determined using memoryless quantization scheme using the codebook $\mathcal{F}_{\mathrm{cb}}$. The quantized CDI for the next block (the $(n+1)$th block) is obtained by informing the transmitter to take a step of size $t_n \in \mathcal{T}$ along the geodesic in the direction of a specified element of $\mathcal{F}_{\mathrm{cb}}$. Accordingly, in Fig. 1, $\mathbf{P}_1$ is obtained by taking a step from $\mathbf{P}_0$ in the direction of $\mathbf{F}_5$, and $\mathbf{P}_2$ is obtained by taking a step from $\mathbf{P}_1$ in the direction of $\mathbf{F}_{10}$.

To formalize the procedure, let us use $g(\mathbf{\Psi})$ to denote the metric by which the receiver evaluates the quality of the unitary matrix $\mathbf{\Psi}$ that represents the CDI. Depending on the scenario, this metric may take one of the forms in (2), (3) or (6).

---

[1]Consistent with the established terminology, we will use the term channel direction information (CDI), even though it is, strictly speaking, channel subspace information.

1) *Initialization*: As in the corresponding memoryless quantization schemes, each receiver selects its initial CDI, $\mathbf{P}_0$, as $\mathbf{P}_0 = \arg\max_{\mathbf{F}_i \in \mathcal{F}_{\mathrm{cb}}} g(\mathbf{F}_i)$, and feeds back the corresponding index. Given that feedback index, the transmitter reconstructs $\mathbf{P}_0$ for that receiver. The transmitter then uses the CDI from all receivers to construct its initial transmit precoder $\mathbf{T}_0$ using the appropriate expression; e.g., in the SU-MIMO case, $\mathbf{T}_0 = \mathbf{P}_0$ and in the MU-MIMO case with ZFBF, $\mathbf{T}_0$ is constructed using the expression in (5).

2) *Incremental updates*: Given its CDI for the nth fading block, $\mathbf{P}_n$, each receiver uses (8) to construct its tangent $\boldsymbol{\Delta}(\mathbf{P}_n, \mathbf{F}_j)$ for each $\mathbf{F}_j \in \mathcal{F}_{\mathrm{cb}}$. For each tangent, the receiver considers steps along the corresponding geodesic of size $\bar{t}_i \in \mathcal{T}$, $i = 1, \ldots, L_{\mathrm{step}}$. Using (9), this process generates $L_{\mathrm{cb}} L_{\mathrm{step}} = 2^B$ candidate CDIs, which we denote by $\mathbf{G}_i \in \mathcal{G}_n$. The receiver then determines

$$\mathbf{P}_{n+1} = \arg\max_{\mathbf{G}_i \in \mathcal{G}_n} g(\mathbf{G}_i), \tag{10}$$

and feeds back the indices that enable the transmitter to construct $\boldsymbol{P}_{n+1}$. Those indices are the index of the direction $\mathbf{F}_i \in \mathcal{F}_{\mathrm{cb}}$ and the step size $\bar{t}_i \in \mathcal{T}$ that generated $\boldsymbol{P}_{n+1}$ in (10). The transmitter reconstructs the CDI from each receiver and computes its next transmit precoder $\mathbf{T}_{n+1}$ using the appropriate expression.

Although the basic principle of the incremental approach is based on a rather simple concept, it has a number of interesting properties.

1) The codebook $\mathcal{F}_{\mathrm{cb}}$ that is used for the directions of the incremental update can be chosen to be a subset of the codebook $\mathcal{F}$ that is used in the memoryless quantization process in the initialization step. As such, there is no additional storage requirement. This enables a system designer to take advantage of existing codebooks for memoryless quantization (e.g., [9], [13], [16], [21]), including codebooks with a subset structure; e.g., [38]. Doing so alleviates some of the difficulties associated with constructing codebooks for increments; e.g., [23].

2) Since the proposed technique is based on updates along geodesics, the precoder in (10) lies on the manifold. As a result, the additional projection step employed in the method in [25] is not required.

3) Given its ability to interpolate between points in the codebook for memoryless feedback (cf., Fig. 1), the performance of the proposed technique is less sensitive to the quality of the codebook than conventional memoryless feedback schemes. Indeed, if the channel is constant, this interpolation yields a feedback scheme with higher resolution than the underlying memoryless scheme.

In the following two sections we will develop two schemes that tailor the basic principles described above to different scenarios:

1) The first scheme is a variant of the generic scheme that is specialized for a first-order Gauss-Markov model for the temporal correlation of the channel. The temporal correlation coefficient of the channel is assumed to be known to the transmitter and the receiver and is assumed to be constant. In this *model-based scheme*, the step size at

channel use n can be adapted to the temporal correlation of the channel, and hence all the feedback bits can be assigned to the direction along the geodesic; i.e., $B_{\mathrm{step}} = 0$. This scheme provides substantial performance gains over memoryless schemes when the temporal correlation of the channel is significant. However, like some other existing schemes, the performance degrades if the temporal correlation of the channel is only coarsely estimated, or if the temporal correlation changes significantly.

2) If the temporal correlation of the channel is not accurately known, or if the nature of the environment and the relative motion of the transmitter and the receivers mean that the temporal correlation coefficient may change, a fraction of the feedback budget should be reserved for the size of the geodesic step. This enables the step size to be adapted to the channel realization rather than the channel statistics. By doing so we obtain robustness to the uncertainty in the temporal channel statistics estimation or abrupt changes in the channel conditions. This is in contrast to the first scheme and the methods in [23]–[25], where the corresponding notions of step size are adapted to the channel statistics. Further, the quantization of the step size can be chosen so that this scheme exhibits an interesting *self-reseting* feature; see Section VI.

V. MODEL-BASED INCREMENTAL FEEDBACK SCHEME

In this section, we present a model-based incremental feedback scheme on the Grassmannian manifold where the step size update for each block is determined according to the temporal correlation of the channel. In this scheme, all the feedback bits B are used to represent the direction of the geodesic, i.e., $B_{\mathrm{cb}} = B$ and $B_{\mathrm{step}} = 0$, and the step size is computed at each channel use based on a statistical analysis that we will provide below. We will assume that the evolution of the channel matrix for a generic user k, $\mathbf{H}_{k,n}$, is modelled by an independent first-order Gauss-Markov process. For notational convenience, we will drop the user index k and model the channel at the nth block as

$$\mathbf{H}_n = \beta \mathbf{H}_{n-1} + \sqrt{1 - \beta^2}\, \boldsymbol{\Theta}_n \tag{11}$$

where $\boldsymbol{\Theta}_n$ has independent entries distributed according to $\mathcal{CN}(0, 1)$. The temporal channel correlation coefficient β is modeled by Jakes model according to $\beta = J_0(2\pi f_{\mathrm{d}})$, where $J_0(\cdot)$ is the zeroth order Bessel function and f_{d} is the normalized Doppler frequency. We will denote the singular value decomposition of $\mathbf{H}_n = \mathbf{U}_n \boldsymbol{\Sigma}_n \mathbf{V}_n^H$ and that of $\boldsymbol{\Theta}_n = \mathbf{Q}_n \boldsymbol{\Lambda}_n \mathbf{S}_n^H$, where the elements of the diagonal matrices $\boldsymbol{\Sigma}_n$ and $\boldsymbol{\Lambda}_n$ are arranged in descending order. Generally, we are interested in tracking the dominant subspace of a particular dimension. In the SU-MIMO case that dimension is Q, for the MU-MISO case it is 1 and for the MU-MIMO case it is Q_k. In order to keep the notation generic, we will let M denote the dimension. Accordingly, we let $\bar{\boldsymbol{\Sigma}}_n$ denote the upper-left $M \times M$ block of $\boldsymbol{\Sigma}_n$, and let $\bar{\mathbf{V}}_n$ denote the first M columns of $\mathbf{V}_n$.

To develop a methodology for choosing the step size to be used at the nth incremental step, we denote the previous quantized version of the CDI by $\mathbf{P}_{n-1}$ and observe that we would

like to move from $\mathbf{P}_{n-1}$ to $\bar{\mathbf{V}}_n$, the current (true) CDI. We do not necessarily move precisely in the direction of $\bar{\mathbf{V}}_n$, as the geodesic is specified by a point in the Grassmannian codebook, but, nevertheless, a reasonable guide for the choice of the step size can be obtained by looking at the average value of the distance between $\mathbf{P}_{n-1}$ and $\bar{\mathbf{V}}_n$. Unfortunately, the analysis of the average geodesic distance between $\mathbf{P}_{n-1}$ and $\bar{\mathbf{V}}_n$ is quite involved. Therefore, we will seek insight into that distance by examining the average chordal distance between $\mathbf{P}_{n-1}$ and $\bar{\mathbf{V}}_n$,

$$d_{\mathrm{ch}}(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_n) = \sqrt{M - \left\| \bar{\mathbf{V}}_n^H \mathbf{P}_{n-1} \right\|_F^2}, \qquad (12)$$

and then making the observation that for small geodesic steps the step size can be approximated by a linear function of the chordal distance. That is, given a point on the manifold $\mathbf{F}_i$ and another point $\mathbf{F}(t)$ of the form in (9) we make the approximation

$$d_{\mathrm{ch}}\left(\mathbf{F}_i, \mathbf{F}(t)\right) \approx \gamma t, \qquad (13)$$

where the choice of γ depends on the intended range of the approximation. As explained in Appendix I, it can be shown that for the model in (11) with M chosen to be M_r, the expected value of the chordal distance between $\mathbf{P}_{n-1}$ and $\bar{\mathbf{V}}_n$ is

$$\begin{aligned} \mathrm{E}\left\{ d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_n) \right\} &= \beta^2 \mathrm{E}\left\{ d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1}) \right\} \\ &\quad + (1 - \beta^2) \mathrm{E}\left\{ d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{S}_n) \right\}, \end{aligned} \qquad (14)$$

where the expectation is taken over the channel model in (11), an i.i.d. Gaussian model for $\mathbf{H}_0$, and random codebooks generated according to the isotropic distribution on the manifold.

Using the above analysis, an appropriate choice for the step size for the nth feedback interval would be to choose t_n so that $(\gamma t_n)^2 = \mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_n)\}$. We now seek an analytic expression for such a t_n, or an approximation thereof.

Since $\mathbf{S}_n$ is isotropically distributed on $\mathbb{G}_{M_t, M}$, and is independent of $\mathbf{P}_{n-1}$, it can be shown that [25]

$$\mathrm{E}\left\{ d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{S}_n) \right\} = \frac{M(M_t - M)}{M_t}. \qquad (15)$$

Furthermore, for the memoryless quantization at step $n = 0$, we have that [39]

$$\begin{aligned} &\mathrm{E}\left\{ d_{\mathrm{ch}}^2(\mathbf{P}_0, \bar{\mathbf{V}}_0) \right\} \\ &\approx \frac{1}{2} \left(\frac{M(M_t - M)}{M(M_t - M) + 1} + \frac{\Gamma\left(\frac{1}{M(M_t - M)}\right)}{M(M_t - M) + 1} \right) \\ &\quad \times \left(\mathcal{C}_{M_t, M} 2^B \right)^{\frac{-1}{M(M_t - M)}}, \end{aligned} \qquad (16)$$

where

$$\mathcal{C}_{M_t, M} = \begin{cases} \frac{1}{(M(M_t - M))!} \Pi_{i=1}^M \frac{(M_t - i)!}{(M - i)!} & 0 < M \leq \frac{M_t}{2} \\ \frac{1}{(M(M_t - M))!} \Pi_{i=1}^{M_t - M} \frac{(M_t - i)!}{(M_t - M - i)!} & \frac{M_t}{2} < M \leq M_t \end{cases} \qquad (17)$$

and $\Gamma(\cdot)$ denotes the gamma function. Using these expressions, we satisfy $(\gamma t_1)^2 = \mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_0, \bar{\mathbf{V}}_1)\}$ if we choose

$$t_1 = \frac{1}{\gamma} \sqrt{\gamma^2 \beta^2 \xi_0^2 + \frac{(1 - \beta^2) M(M_t - M)}{M_t}}, \qquad (18)$$
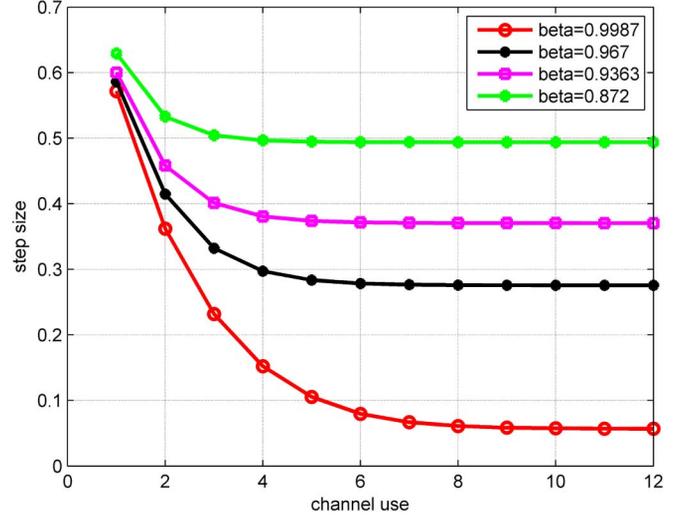
Fig. 2. The evolution of the step size with respect to the channel use for various values of β.

where ξ_0^2 denotes $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_0, \bar{\mathbf{V}}_0)\}$ in (16).

Given the incremental nature of the proposed scheme, $\mathbf{P}_{n-1}$ is correlated with $\mathbf{P}_{n-2}$ and this significantly complicates the evaluation of $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1})\}$ for $n > 1$. However, as shown in Appendix II, by (i) modelling the region that can be spanned by a step of size t_{n-1} from $\mathbf{P}_{n-2}$ by a spherical cap; (ii) partitioning that cap according to the spherical cap approximation of the Voronoi cell of each element of the underlying Grassmannian codebook; and (iii) analyzing the relative volumes of those caps, we can obtain a recursive approximation for $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1})\}$. Using that approximation we obtain the following recursive approximation for t_n for $n \geq 2$ which ensures that $(\gamma t_n)^2 \approx \mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_n)\}$

$$t_n = \frac{1}{\gamma} \sqrt{\mu \gamma^2 \beta^2 t_{n-1}^2 2^{\frac{-B}{M(M_t - M)}} + \frac{(1 - \beta^2) M(M_t - M)}{M_t}} \qquad (19)$$

where μ is a correction factor due to the Voronoi region approximation and is dependent on the number of feedback bits assigned to the codebook.

The evolution of the step size for several values of β is plotted in Fig. 2. From that figure it can be seen that during the first few channel uses, the step size value is large in order to mitigate the quantization errors and the channel evolution. As the number of channel uses increases, the quantization error decreases and the step size converges to a steady state value (i.e., a fixed point) dependent on β. Using (19), that fixed point, which we denote by t_∞, can be found to be

$$t_\infty(\beta) = \sqrt{\frac{(1 - \beta^2) M(M_t - M)}{M_t \gamma^2 \left(1 - \mu \beta^2 2^{\frac{-B}{M(M_t - M)}}\right)}}. \qquad (20)$$

An important feature of the proposed incremental feedback scheme is its intrinsic ability to recover, autonomously, from feedback errors. This feature is illustrated in Fig. 3. In that figure, the points $\mathbf{P}_{un,n}$ denote the sequence of CDI's that would be chosen in case of unlimited feedback, while the point $\mathbf{P}_n$ denotes the quantized version of the CDI for block n under
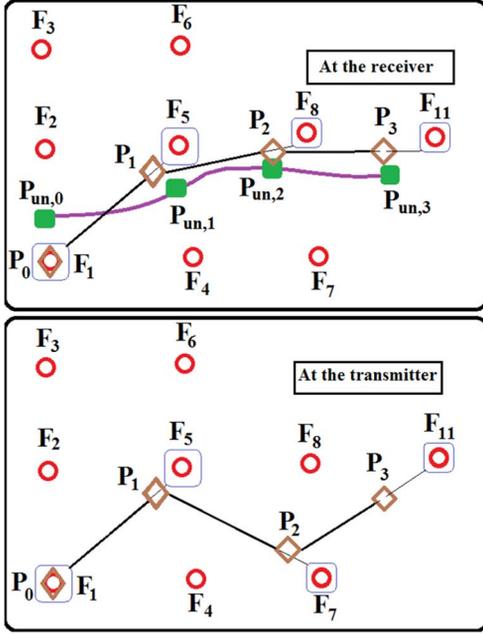
Fig. 3. A pictorial representation of how the incremental schemes can mitigate error propagation in case of feedback errors.

the limited feedback model and the points $\{\mathbf{F}_i\}$ denote the elements of the conventional Grassmannian codebook $\mathcal{F}_{\mathrm{cb}}$. The initial quantized CDI $\mathbf{P}_0$ is determined using memoryless quantization scheme using the codebook $\mathcal{F}_{\mathrm{cb}}$. Now, let's assume that there were no feedback errors during the initialization phase nor in the first update ($n = 0$ and $n = 1$). As a result, at channel use $n = 1$, both the transmitter and receiver have the same $\mathbf{P}_1$. Now consider the case that at the second channel use, the receiver decides to take a step in the direction of $\mathbf{F}_8$ and that errors in feeding back the indices lead the transmitter to interpret the direction of the geodesic as being that associated with $\mathbf{F}_7$. As can be seen by comparing the upper and the lower parts of Fig. 3 the transmitter and the receiver assume totally different values for the quantized CDI. Now, in the third channel use, the receiver chooses to take a step of size t_3 in the direction of $\mathbf{F}_{11}$. If there is no feedback error, the transmitter will move from the previous point $\mathbf{P}_2$ (which, due to the feedback error, is in the wrong place) with a step size t_3 in the direction of $\mathbf{F}_{11}$. Even though the points $\mathbf{P}_2$ at the transmitter and the receiver are not the same, the new points $\mathbf{P}_3$ become closer to each other, and they will get closer still with successive channel uses. The main reason behind this property is that the incremental scheme moves from the current point to a *fixed* point determined by an element of the codebook $\mathcal{F}_{\mathrm{cb}}$.

As we will show in the simulation section, the above model-based incremental technique can achieve performance gains in highly correlated channel and provides better results, under a variety of channel conditions, than other similar schemes that employ some sort of geodesic interpolation.

VI. A MORE ROBUST INCREMENTAL FEEDBACK SCHEME

In the previous section, we developed a model-based incremental scheme, in which the step size is a function of the temporal correlation of the channel β. Although that scheme en-

ables us to retain the full resolution of the memoryless codebook for the directions of the updates, it is inherently sensitive to the accuracy of the channel model. In this section we will describe how the quantization scheme for the step size in the generic form of our approach in Section IV can be chosen so as to develop a more robust incremental scheme.

Once the bit budget B has been partitioned into components for the directions, B_{cb}, and the step size, B_{step}, of the geodesic updates, there are a variety of ways in which the scalar quantization scheme $\mathcal{T} = \{\bar{t}_i\}$ on $[0,1]$ for the step size t can be chosen. A natural, if somewhat generic, approach would be to consider the Lloyd algorithm; e.g., [40] and references therein. If the temporal correlation of the channel model, β, is known and does not change in time, then obtaining a good quantization of $[0,1]$ is quite straightforward. However, in practice the temporal correlation is likely to vary significantly in time, and a more sophisticated approach would be to postulate a distribution for β, and to apply Lloyd's algorithm to the resulting family of models.

Having said that, as we now explain, in our application a case can be made for choosing $\bar{t}_{L_{\mathrm{step}}} = 1$. (Recall that $L_{\mathrm{step}} = 2^{B_{\mathrm{step}}}$.) First, we observe that when the channel is uncorrelated in time, the representatives $\bar{\mathbf{V}}_n$ (cf. (11)) are isotropically distributed on the manifold. Since good codebooks $\mathcal{F}_{\mathrm{cb}}$ are approximately isotropically distributed, if the channel is likely to change rapidly, at least at some points in time, including a step size of 1 in $\mathcal{T}$ is a natural choice, as it will allow the algorithm to "track" the rapidly varying channel. This choice also imbues the approach with a desirable "self-resetting" property, in the sense that whenever the receiver chooses a step size of 1 (and there are no errors in the feedback at that step), both the transmitter and receiver will move to the same point on the manifold (i.e., the same codeword in the codebook), no matter where they were at the previous step. This can be illustrated using Fig. 3, in which there is a feedback error at the second channel use that results in the transmitter and receiver having different values for $\mathbf{P}_2$. If the step size chosen at step $n = 3$ is $t_3 = 1$ (and no feedback errors occur in this step), then both the transmitter and receiver take a step of size 1 in the direction of $\mathbf{F}_{11}$. Hence, even though they have different values for $\mathbf{P}_2$, both the receiver and the transmitter have $\mathbf{P}_3 = \mathbf{F}_{11}$. As such, in just one step the incremental system has recovered from the feedback error in step $n = 2$. Choosing a step-size quantization that includes $\bar{t}_{L_{\mathrm{step}}} = 1$ also helps to develop insight into the choice of the codebook size. When the receiver chooses a step size of 1, the system behaves like a memoryless system with $B - B_{\mathrm{step}}$ bits; i.e., a memoryless system with $2^{B - B_{\mathrm{step}}}$ codewords. There are a number of guidelines for choosing the size of memoryless codebooks in temporally-uncorrelated, spatially i.i.d. Gaussian channels (e.g., [33]–[35], [39]), and since the proposed robust incremental scheme has $2^{B_{\mathrm{step}}} - 1$ choices for the step size other than 1, it will perform at least as well as a memoryless system with a codebook of $2^{B - B_{\mathrm{step}}}$ codewords.

As the above discussion suggests, the choice of B_{step} involves trade-offs between the performance in highly-correlated channels and the performance in rapidly-varying channels. Given the possibility of significant changes in the channel correlation over time, a reasonable approach would be to

choose $B_{\text{step}} = 1$, and to leave the remaining $B - 1$ bits to index the directions of the update on the manifold. As we will demonstrate in the simulations section, this choice results in a scheme that provides robust performance in the presence of significant changes in the channel correlation.

If we make the choice $B_{\text{step}} = 1$, and if we follow the suggestion above that one of the points in the step-size quantization should be 1, then $\mathcal{T} = \{\bar{t}_1, 1\}$ and the remaining design decision is to choose $\bar{t}_1$. While a (constrained) Lloyd approach could be used, a simple alternative is to consider first-order Gauss-Markov models of the form in (11) and to postulate a distribution for the temporal correlation parameter β based on the scenarios that are expected to be encountered. A reasonable choice for the remaining quantization point for the step size is then

$$\bar{t}_1 = \mathrm{E}_\beta \{ t_\infty(\beta) \}, \tag{21}$$

where $t_\infty(\beta)$ is the fixed point of the step-size iteration for the model-based scheme; cf. (20).

Fig. 4. Performance comparison for a SU-MISO system under a channel model in which $\beta = 0.997$.

VII. SIMULATIONS

In this section, we examine the performance of the proposed model-based and robust incremental schemes in a variety of different communication configurations. In each configuration, the channels are modelled using the first-order Gauss-Markov model in (11), with different, and possibly time-varying, values for the correlation coefficient, β. To provide context for the values of β, we will use the relation $\beta = J_0(2\pi f_d)$ from Jakes' model, and the parameters from the IEEE 802.16m standard [41], in which the carrier frequency is 2.5 GHz and the feedback interval is 5 ms.

We will compare the performance of the proposed schemes against the conventional memoryless quantization scheme, and the differential rotation feedback scheme [25]. In the case of systems based on beamforming, we will also compare against the polar-cap differential feedback scheme [24]. For our model-based and robust incremental schemes, the underlying Grassmannian codebook is designed using the technique in [21]. However, the performance of the proposed schemes is almost indistinguishable at the scale of the figures if that codebook is replaced by the so-called PMI codebook (e.g., [4]) in LTE-Advanced. For the other differential schemes, we would like to thank the authors of these schemes for providing us with their codebooks for the updates. For the initialization of all the differential schemes we will adopt the same Grassmannian codebook, and, therefore, all the schemes start from the same point on the manifold. We assume that the base station is equipped with $M_t = 4$ antennas, while the number of receive antennas will be determined by the communication scheme. The SNR is fixed at 10 dB and the feedback budget is $B = 4$ bits per channel use. For this number of feedback bits, setting the Voronoi correction factor in the model-based scheme so that $0.9 < \mu < 0.95$ yields the best performance.

We first consider the case of a single-user multiple-input single-output (SU-MISO) system in which the transmitter uses the vector indexed by the feedback scheme as the beamforming vector; cf. (2). In Fig. 4, have plotted the evolution of the average achievable rate over 1000 realizations of the channel model in (11) for different feedback schemes as the receiver feeds back information. In the considered scenario, the channel is temporally correlated with $\beta = 0.997$, which corresponds to a velocity of 1 km/h. For the implementation of the proposed robust incremental scheme, we present our results for two different choices for the quantization of the step size, $\mathcal{T}$. In the first case (denoted 'Robust scheme-adp'), the step size quantization is $\mathcal{T}_1 = \{t_n(0.997), 1\}$, where $t_n(0.997)$ is the step size chosen for the model-based method with $\beta = 0.997$; cf. (19). In the second case (denoted 'Robust scheme-fix'), the step size quantization is constant in time and is chosen to be $\mathcal{T}_2 = \{\bar{t}_1, 1\}$, where $\bar{t}_1$ is the average of the fixed points of the model based scheme (cf. (21)) over a uniform distribution of β on [0.930, 0.999]. From Fig. 4, we can see that the model-based incremental scheme provides better performance than the other schemes. The performance of the robust incremental scheme with fixed step size quantization $\mathcal{T}_2$ is close to that of the differential rotation scheme and approximately the same as that of the polar-cap differential scheme, while the robust incremental scheme with time-varying step size quantization, $\mathcal{T}_1$, provides improved performance compared to the differential rotation and polar-cap differential schemes. In the scenario of Fig. 4, all the differential schemes perform significantly better than the 4-bit memoryless scheme, but this scenario represents an idealized case in which the channel follows a first-order Gauss-Markov model with significant temporal correlation, the receiver has perfect knowledge of the temporal correlation, and that correlation remains unchanged during the data transmission period.

In Fig. 5, we consider a scenario in which the channel is, initially, temporally correlated with $\beta = 0.997$, and then, at channel use 7, the channel correlation drops to $\beta = 0.5$. This corresponds to a velocity change from 1 km/h to roughly 20 km/h. The transmitter has no knowledge of this change, and hence the proposed model-based scheme continues to choose a step size of $t_n(0.997)$ even after channel use 7. Similarly the
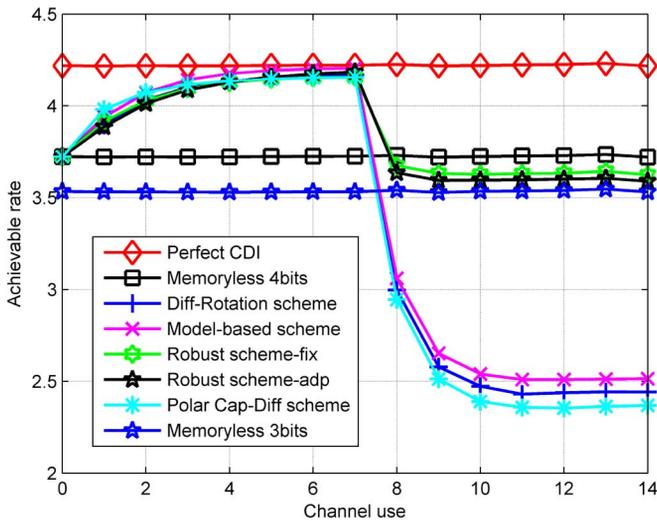
Fig. 5.   Performance comparison for a SU-MISO system under a channel model in which β drops from 0.997 to 0.5 at channel use 7.

Fig. 6.   Performance comparison for a SU-MIMO system under a channel model in which β drops from 0.872 to 0 at channel use 7.

differential schemes continue to use the (small) step sizes associated with $\beta = 0.997$. Prior to the change in the channel correlation, all the schemes provide better performance than the 4-bit memoryless scheme, as the results in Fig. 4 predict. However, their performance is quite different after that change. For this reason, we have added an additional benchmark in the form of a memoryless scheme that uses 3 bits of feedback per channel use. As expected from the discussion in Section VI, the performance of the proposed robust scheme remains better than that of the 3-bit memoryless scheme when the channel correlation changes. (The 3-bit memoryless scheme is equivalent to always choosing a step size of 1 in the proposed robust scheme.) In contrast, the performance of the other differential feedback schemes drops quite a long way below that of the 3-bit memoryless scheme, due to the mismatch between the actual value of β, namely 0.5, and the value of β that was used to choose the step size, namely 0.997. An additional observation from Fig. 5 is that the robust incremental scheme with the fixed step size quantization $\mathcal{T}_2$ has better performance than the robust scheme with the adaptive step-size quantization $\mathcal{T}_1$ when the channel correlation changes, and that it does not require accurate estimation of the channel temporal coefficient β. Accordingly, in the following simulations, we will only present results for the robust scheme with the fixed step-size quantization, $\mathcal{T}_2$.

In Fig. 6 we consider a SU-MIMO system in which the receiver has two antennas and the (four-antenna) transmitter sends two symbols per channel use. The transmitter performs linear precoding with equal power loading using the matrix specified by the feedback scheme; cf. (2). The channel follows the Gauss-Markov model in (11), with the channel correlation changing from $\beta = 0.872$ to $\beta = 0$ at channel use 7. That is, the channel suddenly becomes uncorrelated. (For the IEEE 802.16m parameters that we consider, a correlation of $\beta = 0.872$ corresponds to a speed of 10 km/h.) Even though the channel becomes completely uncorrelated, the proposed robust incremental scheme still provides better performance than the 3-bit memoryless scheme. It also provides significantly better performance than the other competing schemes when the
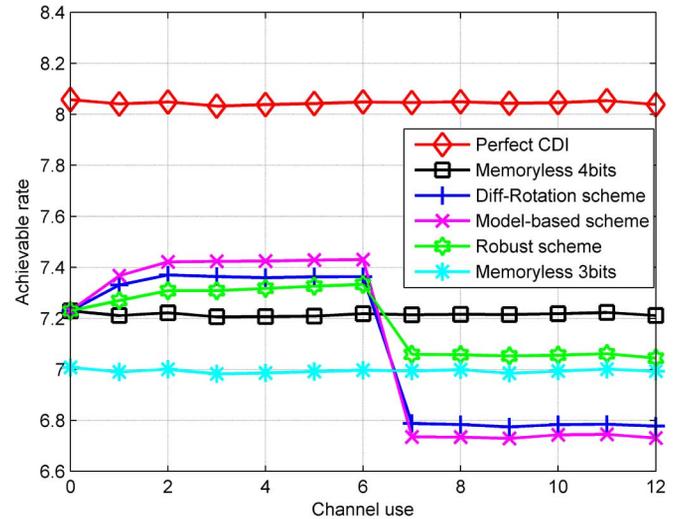
channel becomes uncorrelated. (The fact that the gain of all the schemes in the preliminary stage in Fig. 6 is less than the gain in Fig. 5 is due, in large part, to the lower level of correlation, i.e., the smaller value of β.)

In Figs. 7 and 8, we examine the performance of the considered feedback schemes in the above four-input two-output SU-MIMO system in cases in which the channel correlation β is constant, but is under- or over-estimated, respectively. In Fig. 7 the actual channel correlation is $\beta = 0.997$, whereas it is estimated to be 0.872. In Fig. 8, the actual correlation is $\beta = 0.872$, but it is estimated as being 0.997. Figs. 7 and 8 demonstrate the sensitivities of the existing techniques and the sensitivity of the proposed model-based technique to misestimation of the channel correlation coefficient. The sensitivity arises from the resulting misadaptation of the step size in these techniques. The sensitivity is greater in the case of over-estimation of the correlation coefficient, because that results in small step sizes that are unable to track the variation of the channel. Figs. 7 and 8 also demonstrate the manner in which the proposed robust scheme overcomes the sensitivity to misestimation of the correlation coefficient. In both cases, the proposed robust scheme provides significantly better performance than the other competing schemes. In the case of over-estimation in Fig. 8, it provides better performance than the 4-bit memoryless scheme, whereas the performance of the other competing schemes falls significantly below that of the 3-bit memoryless scheme.

We now examine the performance of the considered feedback schemes in MU-MISO and MU-MIMO systems. Fig. 9 presents results for a MU-MISO system with 4 users that employs Zero-Forcing Beamforming (ZFBF) with uniform power loading; cf. (5). The correlation coefficient of the channel is $\beta = 0.936$, which corresponds to a speed of 7 km/h, and is fixed and precisely known. From Fig. 9, it can be seen that the model-based incremental scheme has the best tracking properties among the competing schemes. In many ways, the relative performance of the considered schemes in this scenario is similar to that for the SU-MISO scenario in Fig. 4.

In order to illustrate the performance of the various schemes in a multiuser scenario in which the channel correlation changes,
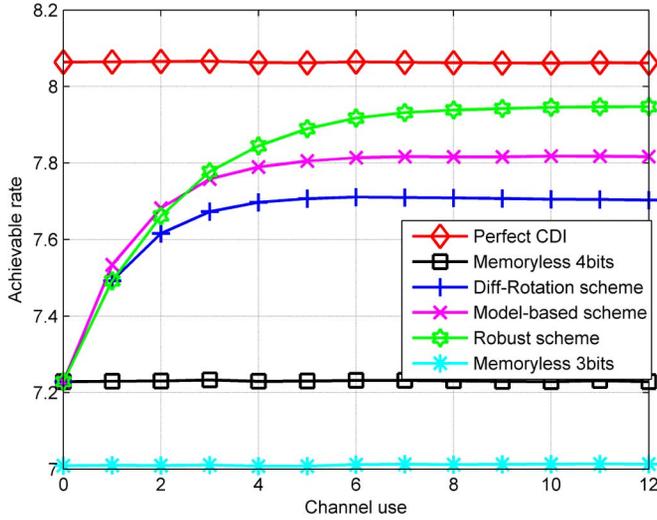
Fig. 7. Performance comparison for a SU-MIMO system under a channel model in which β is constant, but underestimated.
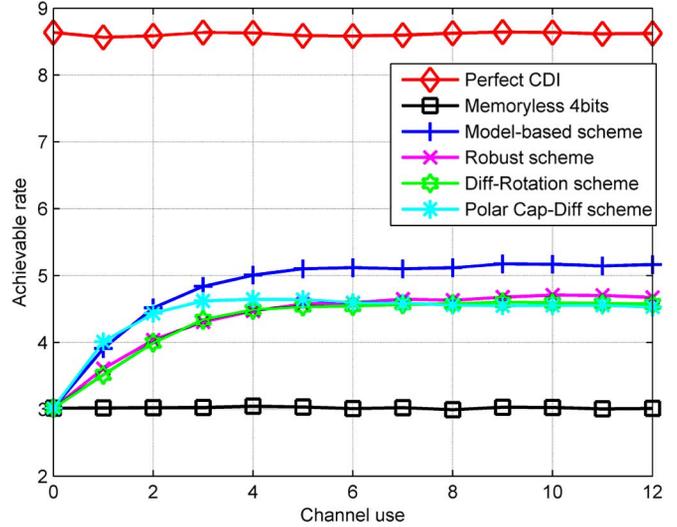
Fig. 9. Performance comparison for a MU-MISO system that employs ZFBF under a channel model with $\beta = 0.936$.
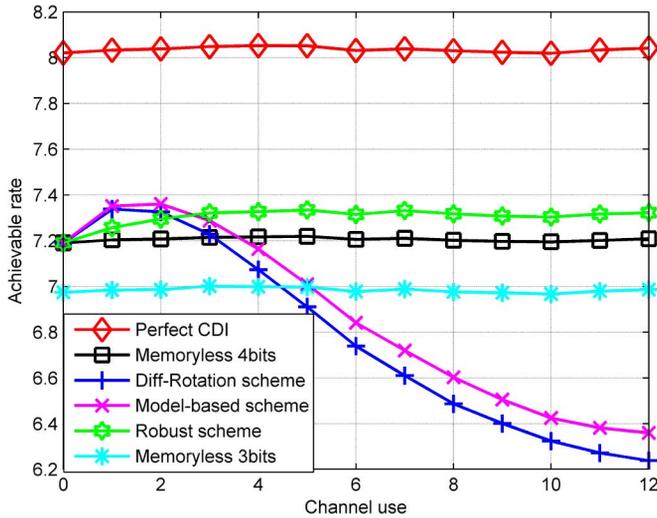
Fig. 8. Performance comparison for a SU-MIMO system under a channel model in which β is constant, but overestimated.
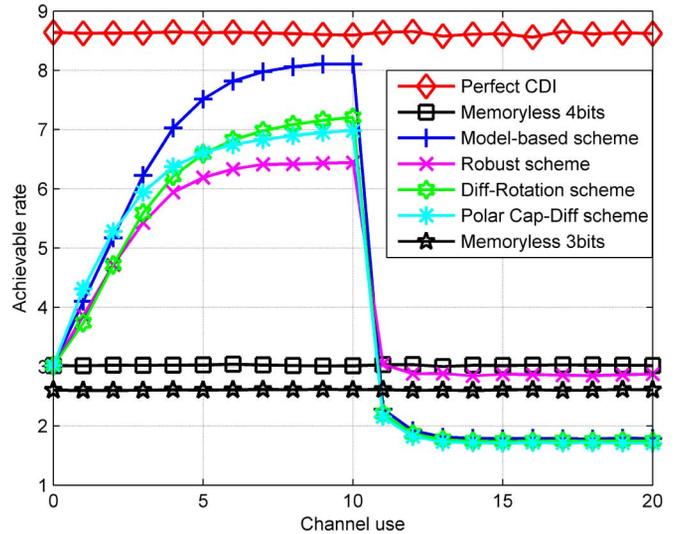
Fig. 10. Performance comparison of MU-MIMO systems with ZFBF and varying β.

in Fig. 10 we consider the performance of the 4-user MU-MISO system that employs ZFBF in an environment in which the channels are initially temporally correlated with $\beta = 0.997$ and then that correlation drops to 0.5 at channel use 11. As can be seen from the figure, initially all the competing methods provide better performance than the 4-bit memoryless scheme, but when the channel correlation drops, only the proposed robust scheme is able to maintain better performance than the 3-bit memoryless scheme.

In Fig. 11, we consider a 2-user MU-MIMO system in which each receiver has two antennas and the transmitter sends two symbols to each receiver at each channel use, using block-diagonalization (BD) with equal power loading; cf. (7). The channels are initially temporally correlated with $\beta = 0.997$ and then that correlation drops to 0.872 at channel use 11. In this scenario, the proposed robust scheme maintains performance better than the 4-bit memoryless scheme, even after the drop in correlation,

whereas the performance of the other competing schemes drops below that of the 3-bit memoryless scheme.

One of the inherent weaknesses in incremental quantization schemes is their sensitivity to errors, and, in particular, the propagation of the effects of those errors. In Fig. 12, we examine the performance of the considered schemes in the presence of a feedback error, and in Fig. 13 we examine the performance in the presence of both a feedback error and errors in the estimation of the correlation coefficient, β. For simplicity, we consider the case of an SU-MISO system. In the scenario considered in Fig. 12, the correlation coefficient is $\beta = 0.936$, and a single feedback error occurs at the 6th channel use. The error is modelled as a switch to another element of the codebook, with each other element having equal probability, $1/(2^{B_{cb}} - 1)$. The figure shows how both the proposed model-based scheme and the proposed robust scheme are able to recover from the feedback error and return to the performance level that was achieved
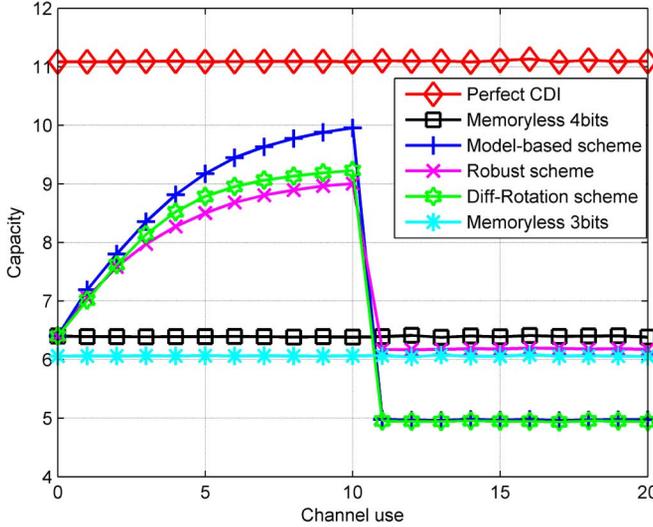
Fig. 11.   Performance comparison of MU-MIMO systems with BD and varying β.
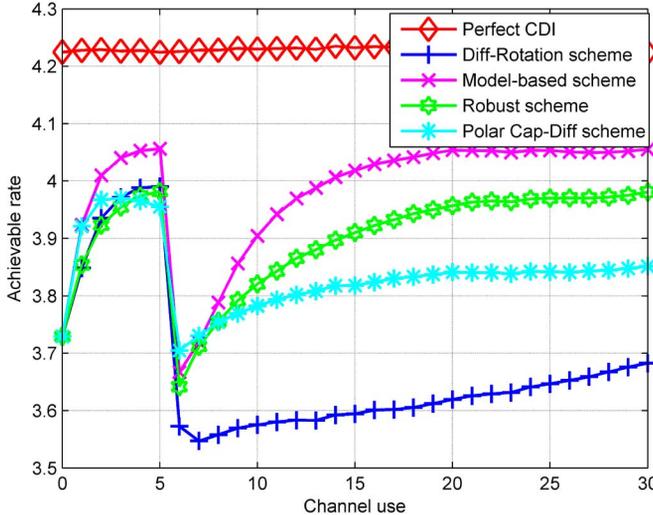
Fig. 12.   Performance comparison for a SU-MISO system under a channel model with $\beta = 0.936$ in the presence of a feedback error at channel use 6.

Fig. 13.   Performance comparison for a SU-MISO system under a channel model in which β is constant, but over-estimated, in the presence of a feedback error at channel use 6.

prior to the error, whereas for the polar cap and differential rotation schemes, the error has a lasting impact that would require an external resetting mechanism to resolve. To explore the combined impact of a feedback error and error in the estimation of the correlation coefficient, we repeated the previous experiment for the case in which the actual correlation coefficient is $\beta = 0.936$, but it is estimated as being 0.967. This corresponds to an error of about 3%. The results for this case are provided in Fig. 13. In this scenario, the proposed robust scheme recovers more quickly than the proposed model-based scheme. Perhaps more importantly, both are able to recover from the feedback error, whereas for the existing competing schemes the presence of the estimation error in the correlation coefficient appears to hinder the ability to recover from the feedback error.

VIII. CONCLUSION

In this paper, we proposed two implementations of an incremental feedback method for temporally correlated channels. The first scheme is adapted to a first-order Gauss-Markov

channel model and uses all the available resources for direction feedback, while the second scheme divides the resources between direction information and step size feedback. From our simulations, we can conclude that when the channel correlation coefficient β is perfectly estimated at the receiver side, the model-based incremental scheme provides the best performance when compared to other existing schemes. However, if the receiver does not have enough time or resources to estimate the long-term channel statistics with sufficient accuracy, or if those statistics change often, we suggest using the robust incremental scheme. A distinguishing feature of both of the proposed schemes is their intrinsic ability to recover from feedback errors and return to their steady state performance levels.

APPENDIX I

DERIVATION OF (14)

We begin with the channel model in (11), $\mathbf{H}_n = \beta\mathbf{H}_{n-1} + \sqrt{1-\beta^2}\,\boldsymbol{\Theta}_n$, and the singular value decompositions $\mathbf{H}_n = \mathbf{U}_n\boldsymbol{\Sigma}_n\mathbf{V}_n^H$ and $\boldsymbol{\Theta}_n = \mathbf{Q}_n\boldsymbol{\Lambda}_n\mathbf{S}_n^H$. Using the analysis in the Appendix of [25], we find that when M is chosen to be equal to M_r,

$$\mathrm{E}\left\{\|\mathbf{H}_n\mathbf{P}_{n-1}\|_F^2\right\}$$
$$= \mathrm{E}\left\{\beta^2\|\mathbf{H}_{n-1}\mathbf{P}_{n-1}\|_F^2 + (1-\beta^2)\|\boldsymbol{\Theta}_n\mathbf{P}_{n-1}\|_F^2\right\} \quad (22\mathrm{a})$$
$$= \mathrm{E}\left\{\beta^2\|\boldsymbol{\Sigma}_{n-1}\mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\|_F^2 + (1-\beta^2)\|\boldsymbol{\Lambda}_n\mathbf{S}_n^H\mathbf{P}_{n-1}\|_F^2\right\}. \quad (22\mathrm{b})$$

Furthermore, using the unitary invariance of the Frobenius norm we have that

$$\mathrm{E}\left\{\|\mathbf{H}_n\mathbf{V}_n\|_F^2\right\} = \mathrm{E}\left\{\|\mathbf{U}_n\boldsymbol{\Sigma}_n\mathbf{V}_n^H\mathbf{V}_n\|_F^2\right\} = \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_n^2\right)\right\} \quad (23)$$

Using the fact that [25]

$$\mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_n^2\right)\right\} = \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_{n-1}^2\right)\right\} = \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Lambda}_n^2\right)\right\}, \quad (24)$$

we can then deduce that

$$
\begin{aligned}
&\mathrm{E}\left\{\|\mathbf{H}_n\mathbf{V}_n\|_F^2\right\} - \mathrm{E}\left\{\|\mathbf{H}_n\mathbf{P}_{n-1}\|_F^2\right\}\\
&= \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_n^2\right)\right\} - \mathrm{E}\left\{\beta^2\|\boldsymbol{\Sigma}_{n-1}\mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\|_F^2\right.\\
&\qquad\qquad\left. + (1-\beta^2)\|\boldsymbol{\Lambda}_n\mathbf{S}_n^H\mathbf{P}_{n-1}\|_F^2\right\}
\end{aligned}
$$
(25a)

$$
\begin{aligned}
&= \mathrm{E}\left\{\mathrm{tr}\left(\beta^2\boldsymbol{\Sigma}_{n-1}^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\right.\\
&\qquad\left.\left. + (1-\beta^2)\boldsymbol{\Lambda}_n^2\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\right)\right\}.
\end{aligned}
$$
(25b)

Using the properties of the trace operator and the fact that $\|\mathbf{X}\|_F^2 = \mathrm{tr}(\mathbf{X}^H\mathbf{X})$, it can be shown that

$$
\begin{aligned}
&\mathrm{E}\left\{\|\mathbf{H}_n\mathbf{V}_n\|_F^2\right\} - \mathrm{E}\left\{\|\mathbf{H}_n\mathbf{P}_{n-1}\|_F^2\right\}\\
&= \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_n^2\right)\right\} - \mathrm{E}\left\{\|\boldsymbol{\Sigma}_n\mathbf{V}_n^H\mathbf{P}_{n-1}\|_F^2\right\}
\end{aligned}
$$
(26a)

$$
= \mathrm{E}\left\{\mathrm{tr}\left(\boldsymbol{\Sigma}_n^2\left(\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right)\right)\right\}.
$$
(26b)

Using the fact that the expectation and the trace operations are linear, by equating (25b) and (26b), we have that

$$
\begin{aligned}
&\mathrm{tr}\left(\mathrm{E}\left\{\boldsymbol{\Sigma}_n^2\left(\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right)\right\}\right)\\
&= \mathrm{tr}\left(\mathrm{E}\left\{\beta^2\boldsymbol{\Sigma}_{n-1}^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\right.\\
&\qquad\left.\left. + (1-\beta^2)\boldsymbol{\Lambda}_n^2\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\right\}\right). (27)
\end{aligned}
$$

Furthermore, using the fact that the channel gains and directions are independent and by using (24), we find that the relation in (27) can be rewritten as

$$
\begin{aligned}
&\mathrm{tr}\left(\mathrm{E}\left\{\boldsymbol{\Sigma}_n^2\right\}\mathrm{E}\left\{\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right\}\right)\\
&= \mathrm{tr}\left(\mathrm{E}\left\{\boldsymbol{\Sigma}_{n-1}^2\right\}\mathrm{E}\left\{\beta^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right\}\right.\\
&\qquad\left. + (1-\beta^2)\mathrm{E}\left\{\boldsymbol{\Lambda}_n^2\right\}\mathrm{E}\left\{\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right\}\right)
\end{aligned}
$$
(28a)

$$
\begin{aligned}
&= \mathrm{tr}\left(\mathrm{E}\left\{\boldsymbol{\Sigma}_n^2\right\}\mathrm{E}\left\{\beta^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\right.\\
&\qquad\left.\left. + (1-\beta^2)\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\right\}\right). (28b)
\end{aligned}
$$

By defining

$$
\mathbf{A} = \mathrm{E}\left\{\boldsymbol{\Sigma}_n^2\right\}
$$
(29)

and

$$
\begin{aligned}
\mathbf{B} = \mathrm{E}\left\{\beta^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\\
+ (1-\beta^2)\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\\
\left. - \left(\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right)\right\},
\end{aligned}
$$
(30)

and by moving the right hand side of (28) to the left hand side, we can rewrite the relation in (28) as $\mathrm{tr}(\mathbf{AB}) = 0$. Since $\mathbf{A}$ and $\mathbf{B}$ are symmetric, the quantity $\mathrm{tr}(\mathbf{AB})$ can be bounded by [42]

$$
\lambda_{\min}(\mathbf{A})\mathrm{tr}(\mathbf{B}) \le \mathrm{tr}(\mathbf{AB}) \le \lambda_{\max}(\mathbf{A})\mathrm{tr}(\mathbf{B})
$$
(31)

where $\lambda_{\max}(\mathbf{A})$ is the greatest eigenvalue of the matrix $\mathbf{A}$ and $\lambda_{\min}(\mathbf{A})$ is the least eigenvalue. Under channel models in which

$\mathrm{E}\{\boldsymbol{\Sigma}_n^2\}$ has full rank, such as the case of the model in (11), $\lambda_{\min}(\mathbf{A}) > 0$, and therefore the condition $\mathrm{tr}(\mathbf{AB}) = 0$ implies that $\mathrm{tr}(\mathbf{B}) = 0$. Therefore,

$$
\begin{aligned}
\mathrm{tr}(\mathbf{B}) &= \mathrm{tr}\left(\mathrm{E}\left\{\beta^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\right.\\
&\quad + (1-\beta^2)\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\\
&\quad\left.\left. - \left(\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right)\right\}\right) (32a)
\end{aligned}
$$

$$
\begin{aligned}
&= \mathrm{E}\left\{\mathrm{tr}\left(\beta^2\left(\mathbf{I}_M - \mathbf{V}_{n-1}^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_{n-1}\right)\right.\right.\\
&\quad + (1-\beta^2)\left(\mathbf{I}_M - \mathbf{S}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{S}_n\right)\\
&\quad\left.\left. - \left(\mathbf{I}_M - \mathbf{V}_n^H\mathbf{P}_{n-1}\mathbf{P}_{n-1}^H\mathbf{V}_n\right)\right)\right\} (32b)
\end{aligned}
$$

$$
\begin{aligned}
&= \beta^2\mathrm{E}\left\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{V}_{n-1})\right\}\\
&\quad + (1-\beta^2)\mathrm{E}\left\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{S}_n)\right\}\\
&\quad - \mathrm{E}\left\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{V}_n)\right\} = 0 (32c)
\end{aligned}
$$

This concludes the derivation of (14).

For the case in which M is chosen to be less than M_r, the corresponding steps yield an inequality in the form of $\mathrm{tr}(\bar{\mathbf{A}}\bar{\mathbf{B}}) \ge 0$, where $\bar{\mathbf{A}} = \mathrm{E}\{\bar{\boldsymbol{\Sigma}}_n^2\}$ and $\bar{\mathbf{B}}$ takes the form in (32) with $\mathbf{V}_n$ replaced by $\bar{\mathbf{V}}_n$. That implies that $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{V}_n)\} \le \beta^2\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{V}_{n-1})\} + (1-\beta^2)\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \mathbf{S}_n)\}$. Accordingly, the step size in the model-based scheme is designed based on the upper bound on the expected chordal distance between $(\mathbf{P}_{n-1}, \mathbf{V}_n)$

APPENDIX II

APPROXIMATION OF $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1})\}$

In order to find an expression for $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1})\}$, we employ a Voronoi region approximation akin to that in [24]. To do so, we first define the spherical cap centered at $\mathbf{U}$ with radius r to be [39]:

$$
\mathcal{S}_{\mathbf{U}}(r) = \{\mathbf{V} : d_{\mathrm{ch}}(\mathbf{U}, \mathbf{V}) < r; \mathbf{V} \in \mathbb{G}_{M_t, M}\}. (33)
$$

The volume of that spherical cap is:

$$
\mathrm{Vol}\left(\mathcal{S}_{\mathbf{U}}(r)\right) = \mathcal{C}_{M_t, M}r^{2M(M_t - M)}, (34)
$$

where $\mathcal{C}_{M_t, M}$ was defined in (17).

The analysis begins with the point $\mathbf{P}_{n-2}$ on the manifold, and the set of all points on the manifold that can be obtained by taking a step of size t_{n-1} along a geodesic from that point. Using the approximate relationship between the geodesic and chordal distances in (13), we can represent that set of points by the spherical cap $\mathcal{S}_{\mathbf{P}_{n-2}}(\gamma t_{n-1})$. When we update $\mathbf{P}_{n-2}$ to $\mathbf{P}_{n-1}$ we take a step of size t_{n-1} in the direction of one of the 2^B elements of the codebook $\mathcal{F}$. If we model those elements as being isotropically distributed, then we implicitly partition $\mathcal{S}_{\mathbf{P}_{n-2}}(\gamma t_{n-1})$ into 2^B spherical caps of equal radii. If we make the (high-resolution) Voronoi approximation that those 2^B spherical caps cover $\mathcal{S}_{\mathbf{P}_{n-2}}(\gamma t_{n-1})$ without overlap, then each has a volume that is $1/2^B$ of the volume of $\mathcal{S}_{\mathbf{P}_{n-2}}(\gamma t_{n-1})$. Using (34), this means that the radius, r_n of each of the 2^B spherical caps is such that

$$
2^B\mathcal{C}_{M_t, M}r_n^{2M(M_t - M)} = \mathcal{C}_{M_t, M}(\gamma t_{n-1})^{2M(M_t - M)}. (35)
$$

That is,

$$r_n^2 = (\gamma t_{n-1})^2 2^{\frac{-B}{M(M_t-M)}}. \tag{36}$$

Now, when we actually take the step from $\mathbf{P}_{n-2}$ we move in the direction of $\mathbf{P}_{n-1}$, which is the centre of the one of the 2^B spherical caps that contains $\bar{\mathbf{V}}_{n-1}$. That is, $\bar{\mathbf{V}}_{n-1}$ lies in $\mathcal{S}_{\mathbf{P}_{n-1}}(r_n)$. Therefore, up to the accuracy of the above approximations, $\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_{n-1}, \bar{\mathbf{V}}_{n-1})\} = r_n^2 = (\gamma t_{n-1})^2 2^{\frac{-B}{M(M_t-M)}}$. In order to account for the errors incurred in those approximations, and in particular the errors incurred in the Voronoi approximation, we apply a correction factor $\mu \in [0, 1]$ to that result and refine our approximation to

$$\mathrm{E}\{d_{\mathrm{ch}}^2(\mathbf{P}_n, \bar{\mathbf{V}}_n)\} \approx \mu(\gamma t_n)^2 2^{\frac{-B}{M(M_t-M)}}. \tag{37}$$

Although it appears to be difficult to obtain an analytic expression for μ, it depends on the number of codewords in the codebook and the dimension of the manifold and can be determined, off-line, using straightforward numerical techniques. In our simulations, we found that for a codebook of 2^4 codewords, setting $0.9 < \mu < 0.95$ gives good performance. As the number of feedback bits increases, the accuracy of the Voronoi region approximation also increases, and hence μ approaches 1.

ACKNOWLEDGMENT

REFERENCES

[1] G. Caire and K. Kumar, "Information theoretic foundations of adaptive coded modulation," *Proc. IEEE*, vol. 95, no. 12, pp. 2274–2298, 2007.

[2] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, Sep. 2007.

[3] Q. Li, G. Li, W. Lee, M. Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO techniques in WiMAX and LTE: A feature overview," *IEEE Commun. Mag.*, vol. 48, no. 5, pp. 86–92, May 2010.

[4] J. Lee, J.-K. Han, and J. Zhang, "MIMO technologies in 3GPP LTE and LTE-Advanced," *EURASIP J. Wirel. Commun. Netw.*, vol. 2009, pp. 3:1–3:10, Mar. 2009.

[5] C. Lim, T. Yoo, B. Clerckx, B. Lee, and B. Shim, "Recent trend of multiuser MIMO in LTE-Advanced," *IEEE Commun. Mag.*, vol. 51, no. 3, pp. 127–135, Mar. 2013.

[6] P. de Kerret and D. Gesbert, "CSI sharing strategies for transmitter cooperation in wireless networks," *IEEE Wireless Commun. Mag.*, vol. 20, no. 1, pp. 43–49, Feb. 2013.

[7] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[8] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer, 1991.

[9] D. J. Love, R. W. Heath, Jr., V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, Oct. 2008.

[10] U. Salim, D. Gesbert, and D. Slock, "Combining training and quantized feedback in multiantenna reciprocal channels," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1383–1396, 2012.

[11] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.

[12] D. J. Love, R. W. Heath, Jr., and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, Oct. 2003.

[13] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Oct. 2005.

[14] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

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

[16] B. Clerckx, Y. Zhou, and S. Kim, "Practical codebook design for limited feedback spatial multiplexing," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, 2008, pp. 3982–3987.

[17] I. S. Dhillon, R. W. Heath, T. Strohmer, and J. A. Tropp, "Constructing packings in Grassmannian manifolds via alternating projection," *Exper. Math.*, vol. 17, no. 1, pp. 9–35, 2008.

[18] K. Schober, P. Janis, and R. Wichman, "Geodesical codebook design for precoded MIMO systems.," *IEEE Commun. Lett.*, vol. 13, no. 10, pp. 773–775, 2009.

[19] B. Mondal, T. A. Thomas, and M. Harrison, "Rank-independent codebook design from a quaternary alphabet," in *Conf. Rec. 41st Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, 2007, pp. 297–301.

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

[21] A. Medra and T. N. Davidson, "Flexible codebook design for limited feedback systems via sequential smooth optimization on the Grassmannian manifold," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1305–1318, March 2014.

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

[23] R. W. Heath, T. Wu, and A. C. K. Soong, "Progressive refinement of beamforming vectors for high-resolution limited feedback," *EURASIP J. Adv. Signal Process.*, vol. 2009, pp. 6:1–6:13, Feb. 2009.

[24] J. Choi, B. Clerckx, N. Lee, and G. Kim, "A new design of polar-cap differential codebook for temporally/spatially correlated MISO channels," *IEEE Trans. Wireless Commun.*, vol. 11, no. 2, pp. 703–711, 2012.

[25] T. Kim, D. Love, and B. Clerckx, "MIMO systems with limited rate differential feedback in slowly varying channels," *IEEE Trans. Commun.*, vol. 59, no. 4, pp. 1175–1189, 2011.

[26] T. Kim, D. Love, B. Clerckx, and S. Kim, "Differential rotation feedback MIMO system for temporally correlated channels," in *Proc. IEEE Global Telecommun. Conf.*, 2008, pp. 1–5.

[27] O. El Ayach and R. Heath, "Grassmannian differential limited feedback for interference alignment," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6481–6494, 2012.

[28] L. Liu and H. Jafarkhani, "Novel transmit beamforming schemes for time-selective fading multiantenna systems," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4767–4781, 2006.

[29] T. Inoue and R. Heath, "Grassmannian predictive coding for limited feedback multiuser MIMO systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2011, pp. 3076–3079.

[30] K. Kim, T. Kim, D. Love, and I. H. Kim, "Differential feedback in codebook-based multiuser MIMO systems in slowly varying channels," *IEEE Trans. Commun.*, vol. 60, no. 2, pp. 578–588, 2012.

[31] T. Kim, D. Love, and B. Clerckx, "Leveraging temporal correlation for limited feedback multiple antennas systems," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2010, pp. 3422–3425.

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

[33] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2006.

[34] N. Jindal, "Antenna combining for the MIMO downlink channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3834–3844, Jul. 2007.

[35] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1473–1482, Oct. 2008.

[36] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.

[37] J. H. Conway, R. H. Hardin, and N. J. A. Sloane, "Packing lines, planes, etc.: Packings in Grassmannian spaces," *Exper. Math.*, vol. 5, pp. 139–159, 1996.

[38] R. H. Gohary and T. N. Davidson, "Non-coherent MIMO communication: Grassmannian constellations and efficient detection," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1176–1205, Mar. 2009.

[39] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, 2008.

[40] R. Gray and D. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.

[41] *IEEE P802.16m-2008 Draft Standard for Local and Metropolitan Area Network*, IEEE Standard 802.16m, 2008.

[42] Y. Fang, K. Loparo, and X. Feng, "Inequalities for the trace of matrix product," *IEEE Trans. Autom. Control*, vol. 39, no. 12, pp. 2489–2490, Dec. 1994.

Ahmed Medra received the B.Sc. degree (with honors) in electrical engineering from Alexandria University, Egypt, in 2005, the M.Sc. degree from Alexandria University, Egypt, in 2010, and the Ph.D degree from McMaster University, Canada, in 2014. Currently, he is a post-doctoral fellow in ECE department at University of Bristish Columbia (UBC), Canada. He had received the Philomathia Foundation Fellowship in 2011 and McMaster International Excellence Award in 2012. His research interest includes MIMO communications, Grassmannian codebook design, linear precoding and Interference Alignment.

Timothy N. Davidson received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He is a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is currently serving as Chair of the Department. Previously, he has served as Acting Director of the School of Computational Engineering and Science for two years, and as Associate Director for three years. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and a 2011 Best Paper Award from the IEEE Signal Processing Society. He has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He has also served as a Guest Co-Editor of issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, and the *EURASIP Journal on Advances in Signal Processing*. He was a General Co-Chair for the 2014 IEEE International Workshop on Signal Processing Advances in Wireless Communications, and a Technical Program Co-Chair for the 2014 IEEE Global Conference on Signal and Information Processing. Dr. Davidson is currently serving as the Past-Chair of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.