Limited Feedback in Massive MIMO Systems: Exploiting Channel Correlations via Noncoherent Trellis-Coded Quantization

Junil Choi*, David J. Love*, and Upamanyu Madhow†
*School of Electrical and Computer Engineering
Purdue University, West Lafayette, IN
†Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA

Abstract—**Accurate channel state information at the transmitter can significantly enhance the performance of multiple antenna systems, but efficient channel quantization techniques must be developed in order to scale such informed transmitter strategies to frequency division duplexed (FDD) massive MIMO systems. Recent results show that noncoherent trellis-coded quantization (NTCQ) is an effective approach for limited feedback transmit beamforming with a large number of transmit antennas. In this paper, we investigate extensions of NTCQ that exploit channel correlations in time and space to significantly reduce the required feedback rate. Our numerical results show that near-optimal beamforming with moderate feedback overhead can be obtained in massive MIMO systems.**

I. Introduction

Multiple-input multiple-output (MIMO) systems have been thoroughly investigated over the last two decades, and the benefits of MIMO in enhancing reliability and throughput are well established. While existing WiFi [1] and cellular standards [2] incorporate MIMO techniques with a moderate number (e.g., up to four or eight) of antennas, there is significant recent interest in "massive MIMO" systems with a significantly larger number (10s-100s) of antennas [3]. It was shown in [4] that, even with very noisy channel estimation, adding more antennas at the base station is always beneficial. Large antenna arrays at base stations can lead to revolutionary increases in the power efficiency (and hence range) [5] and the number of simultaneous users (and hence network capacity) [6] in future cellular systems. We refer to [7] and references therein for further details.

Channel state information (CSI) at the base station can significantly enhance the performance of the massive MIMO, enabling range increases via transmit beamforming and network capacity increases via Space Division Multiple Access (SDMA). Much of the recent literature on massive MIMO assumes that CSI is available via channel reciprocity in time division duplexed (TDD) systems. However, most existing cellular networks are frequency division duplexed (FDD), hence it is of interest to develop massive MIMO upgrades to such systems. A major hurdle in developing CSI for massive MIMO FDD downlinks is the scaling of the overhead required for downlink channel sounding and uplink feedback with the number of antennas. In [8], a novel limited feedback strategy based on noncoherent trellis-coded quantization (NTCQ) is proposed for systems with a large (and even time-varying) number of transmit antennas. NTCQ exploits the duality between optimal beamforming and block noncoherent sequence detection, where the size of the block equals the number of transmit elements. NTCQ can be dynamically adapted to changes in the number of transmit antennas (by adapting the block size) and in the allowable feedback rate (by adapting the underlying coded modulation scheme, including constellation and code rate).

In this paper, we develop advanced limited feedback strategies based on NTCQ which exploit channel temporal and spatial correlations to reduce feedback overhead. Differential codebooks [9]–[16] and adaptive codebooks [17]–[20], which exploit temporal and spatial correlation of channels, respectively, have been studied over the last few years. However, these techniques assume a fixed and moderate number of transmit antennas, and hence do not apply to massive MIMO systems with a large, and potentially variable, number of antennas. The advanced NTCQ techniques proposed in this paper develop differential and spatially adaptive techniques that exploit temporal and spatial correlations, respectively, to improve upon the original NTCQ scheme, while maintaining the flexibility and scalability required for massive MIMO systems.

The remainder of this paper is organized as follows. The limited feedback system model considered in this paper is presented in Section II. We briefly review the original NTCQ scheme [8] in Section III. The proposed advanced NTCQ schemes for temporally and spatially correlated channels are presented in Section IV. Simulation results demonstrating the effectiveness of the proposed schemes are provided in Section V. Section VI contains our conclusions.

II. System Model

We consider the multiple-input single-output (MISO) system, shown in Fig. 1, with n_t transmit antennas at the BS and a single receive antenna transmitting over a frequency flat block fading channel to a single antenna receiver. Among n_t

978-1-4673-5239-0/13/$31.00 ©2013 IEEE

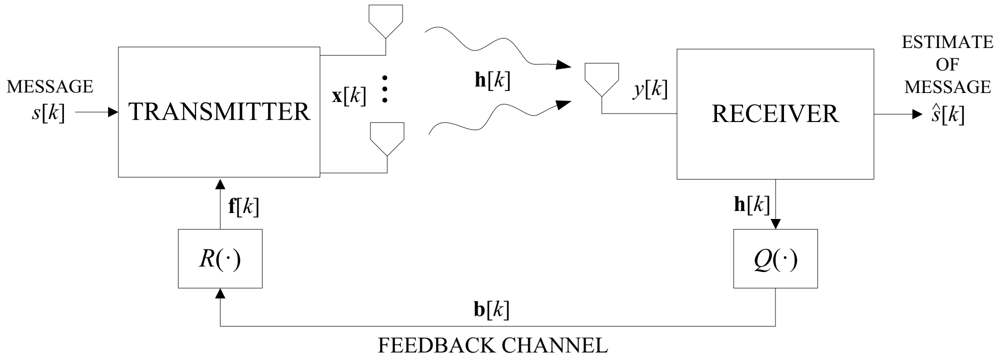

Fig. 1: Multiple-input single-output communications system with feedback channel. The receiver quantizes the current channel realization using NTCQ and sends feedback to the transmitter. The transmitter uses an NTCQ decoder to reconstruct the beamformer.

transmit antennas, we assume only $M(k)$ antennas are used for transmission at time k (i.e., if all antennas are always used $M(k) = n_t$). This is a more general assumption for massive MIMO because it covers the case when all transmitters are used *and* the case when only the transmitters experiencing good channel conditions are used for transmission. Optimizing the effective number of transmit antennas in massive MIMO systems can be an interesting research topic, which is beyond the scope of this paper. The received signal y at time instant k is written as

$$y[k] = \mathbf{h}[k]^H \mathbf{x}[k] + n[k] \tag{1}$$

where $n[k]$ denotes the noise which is drawn from $\mathcal{CN}(0, \sigma^2)$, $\mathbf{h}[k]$ is the $\mathbb{C}^{M(k)\times 1}$ channel vector, and $\mathbf{x}[k]$ is the $\mathbb{C}^{M(k)\times 1}$ transmitted signal vector. The channel vector $\mathbf{h}[k]$ can be temporally or spatially correlated, and the exact expression of $\mathbf{h}[k]$ is given in Section IV-A and IV-B for temporally and spatially correlated channels, respectively. The transmitted signal $\mathbf{x}[k]$ is given by $\mathbf{x}[k] = \mathbf{f}[k]s[k]$ where $\mathbf{f}[k]$ is a beamforming vector with $||\mathbf{f}[k]||_2 = 1$ and $s[k]$ is a message stream such that $E[s[k]] = 0$ and $E[|s[k]|^2] = \rho$. The received SNR in this setup is given by

$$\Gamma(\mathbf{f}[k]) = \frac{E[|\mathbf{h}[k]^H \mathbf{f}[k]s[k]|^2]}{E[|n[k]|^2]} = \frac{\rho}{\sigma^2}|\mathbf{h}[k]^H \mathbf{f}[k]|^2. \tag{2}$$

The receiver in the system is assumed to perfectly know $\mathbf{h}[k]$. As shown in Fig. 1, the receiver represents a quantized version of $\mathbf{h}[k]$ using a quantization function $Q(\cdot)$. The quantization bits are sent over a feedback channel to the transmitter. The transmitter then reconstructs $\mathbf{f}[k]$ using a reconstruction function $R(\cdot)$. The main focus of this paper is to design the quantizer $Q(\cdot)$ and reconstructor $R(\cdot)$ for CSI in Fig. 1 for a large number of transmit antenna scenario. This problem can also be viewed as the quantization by the receiver of the unquantized beamforming vector $\mathbf{f}_{\mathrm{opt}}[k]$ using a codebook $\mathcal{C} = \{\mathbf{c}_1, \ldots, \mathbf{c}_{2^B}\}$ known to both the transmitter and receiver.

III. Noncoherent Trellis-coded Quantization

In this section, we briefly overview noncoherent trellis-coded quantization (NTCQ) for beamformer quantization [8]. Channel quantization based on trellis is also studied in [21]–[23]; however, they adopted either suboptimal Euclidean distance quantization or an ad hoc path metric. Our quantizer is based on leveraging a duality between source and channel coding similar to [24]. In [24], maximum likelihood decoding techniques are directly applied to the channel quantization problem. The proposed NTCQ, however, exploits the modulation and coding technique of trellis-coded modulation (TCM) [25], resulting in much better performance than schemes in [24].

First, consider the noncoherent additive white Gaussian noise, block noncoherent channel

$$\mathbf{y} = e^{j\theta}\mathbf{x} + \mathbf{z}, \tag{3}$$

where $e^{j\theta}$ is an unknown constant phase offset, $\mathbf{y} \in \mathbb{C}^{N\times 1}$ is the received signal, $\mathbf{x} \in \mathbb{C}^{N\times 1}$ is the vector of transmitted symbols and $\mathbf{z} \in \mathbb{C}^{N\times 1}$ is the complex Gaussian noise. Note that, for this example, the unknown channel parameter is the phase, θ (but, in general, it could also have an unknown amplitude). $\mathbf{x}$ could have entries taken from constellations such as phase shift keying (PSK) and quadrature amplitude modulation (QAM). If $\mathbf{x}$ consists of PSK constellation, $\mathbf{x}$ has a constant norm while the norm of $\mathbf{x}$ can vary with QAM constellation. We assume $\mathbf{x}$ is of a constant norm for brief explanation.

As in [24], [26], with the generalized likelihood ratio test (GLRT), the estimate of the transmitted vector $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = \operatorname*{argmin}_{\mathbf{x}\in\mathbb{C}^{N\times 1}} \min_{\theta\in[0,2\pi)} \|\mathbf{y} - e^{j\theta}\mathbf{x}\|_2^2, \tag{4}$$

$$= \operatorname*{argmin}_{\mathbf{x}\in\mathbb{C}^{N\times 1}} \min_{\theta\in[0,2\pi)} \|\mathbf{y}\|_2^2 + \|\mathbf{x}\|_2^2 - 2\,\mathrm{Re}(e^{j\theta}\mathbf{y}^H\mathbf{x}), \tag{5}$$

$$= \operatorname*{argmax}_{\mathbf{x}\in\mathbb{C}^{N\times 1}} |\mathbf{y}^H\mathbf{x}|, \tag{6}$$

where (6) follows from the fact that

$$\operatorname*{argmax}_{\mathbf{x}\in\mathbb{C}^{N\times 1}} |\mathbf{y}^H\mathbf{x}| = \operatorname*{argmin}_{\mathbf{x}\in\mathbb{C}^{N\times 1}} \min_{\theta\in[0,2\pi)} \left\{-\mathrm{Re}(e^{j\theta}\mathbf{y}^H\mathbf{x})\right\}.$$

Similarly, the beamforming vector $\mathbf{f}[k]$ that maximizes beamforming gain in (2) maximizes the projection of the channel vector onto the one-dimensional complex subspace spanned by a codeword and is given by

$$\mathbf{f}[k] = \underset{\mathbf{c}\in\mathcal{C}}{\operatorname{argmax}} \left| \mathbf{h}[k]^H \frac{\mathbf{c}}{||\mathbf{c}||_2} \right|^2. \tag{7}$$

This is equivalent to minimizing the projection of the channel vector orthogonal to this subspace, and finding the optimal codeword $\mathbf{c}_{\mathrm{opt}}$ can be performed by optimizing

$$\min_{\alpha\in\mathbb{R}^+} \min_{\theta\in[0,2\pi)} \min_{\mathbf{c}\in\mathcal{C}} \|\mathbf{h}[k] - \alpha e^{j\theta}\mathbf{c}\|_2^2. \tag{8}$$

When the codeword elements are chosen from a PSK constellation, we can set $\alpha = 1$ (this does not minimize the projection orthogonal to $\mathbf{c}$, but minimizes a closely related cost function). However, when the codeword elements can come from a QAM constellation and the codewords may have unequal norm, then it becomes essential to optimize over α in order to solve the original problem (7). The inner minimization over Euclidean distance in (8) can be performed using the Viterbi algorithm. Instead of searching over $\alpha \in \mathbb{R}^+$ and $\theta \in [0, 2\pi)$, we can closely approximate the optimization using a discrete set, $\alpha \in \mathbb{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_{K_\alpha}\}$ and $\theta \in \Theta = \{\theta_1, \theta_2, \ldots, \theta_{K_\theta}\}$. Thus, the vector search (8) is implemented using $K_\alpha \cdot K_\theta$ instances of the Viterbi algorithm. The dynamic range of the amplitude search (and hence the size of the discrete set $\mathbb{A}$) is reduced by replacing the channel vector $\mathbf{h}[k]$ by its normalized version $\bar{\mathbf{h}}[k] = \frac{\mathbf{h}[k]}{||\mathbf{h}[k]||_2}$.

We need to define a path metric to perform the Viterbi algorithm. To do this, first define $\mathbf{p}_t$ as a partial path up to stage t and $\mathrm{out}(\mathbf{p}_t)$ as a function that outputs the sequence of constellation points that correspond to the path $\mathbf{p}_t$. Then we can define the path metric $m(\cdot)$ as

$$\begin{aligned} m(\mathbf{p}_t, \alpha_k, \theta_k) &= \|\bar{\mathbf{h}}_t - \alpha_k e^{j\theta_k}\,\mathrm{out}(\mathbf{p}_t)\|_2^2 \\ &= m(\mathbf{p}_{t-1}, \alpha_k, \theta_k) + \left|\bar{h}_t - \alpha_k e^{j\theta_k}\,\mathrm{out}\left([p_{t-1}\ p_t]^T\right)\right|^2 \end{aligned} \tag{9}$$

where $\bar{\mathbf{h}}_t$ is the vector truncated to the first t entries of $\bar{\mathbf{h}}$, $\bar{h}_t$ is the t^{th} entry of $\bar{\mathbf{h}}$ and p_t is the t^{th} entry of $\mathbf{p}_t$. Note that minimizing the above path metric will minimize Euclidean distance. With the path metric in (9), each channel entry is quantized by each transition in the trellis. It is important to note that, because of the trellis structure or set partitioning of TCQ, only half of the constellation points are used during quantization in each transition, i.e., when we rely on 8PSK constellation only 2bits or a QPSK constellation point is used to quantize each channel entry. However, there are two sets of QPSK constellation points, and the QPSK set used for quantization depends on the state in trellis. This issue is explained in more detail in [8].

Having chosen the best path for each α_k and θ_k using the Viterbi algorithm, we then choose among these paths the one that achieves the best metric. That is, we optimize the tuple $(\mathbf{p}_{\mathrm{best}}, \alpha_{\mathrm{best}}, \theta_{\mathrm{best}})$ by

$$\min_{\mathbf{p}_T\in\mathbb{P}} \min_{\alpha_k\in\mathbb{A}} \min_{\theta_k\in\Theta} m(\mathbf{p}_T, \alpha_k, \theta_k) \tag{10}$$

where $\mathbb{P}$ is a set of all possible paths and $T = M(k)$, i.e., the number of transmit antennas. The beamforming vector $\mathbf{f}[k]$ can be calculated as

$$\mathbf{c}_{\mathrm{opt}} = \mathrm{out}(\mathbf{p}_{\mathrm{best}}), \quad \mathbf{f}[k] = \frac{\mathbf{c}_{\mathrm{opt}}}{\|\mathbf{c}_{\mathrm{opt}}\|_2}. \tag{11}$$

The search over α and θ only increases the complexity of beamforming quantization, not feedback overhead because the transmitter does not need to known α_{best} and θ_{best} to reconstruct the beamforming vector. However, there do exist an additional feedback information for NTCQ. The transmitter have to know the starting state of $\mathbf{p}_{\mathrm{best}}$, which gives $\log_2 S$ additional feedback overhead where S is the number of states in the trellis.

IV. Exploiting Channel Correlations in NTCQ

A. Differential NTCQ Scheme

Most of the literature regarding differential codebooks use limited feedback utilize temporal correlation and consider a fixed and small number of antennas at the transmitter side [9]–[16]. In massive MIMO systems, however, this may not be the case. Therefore, we propose a new differential feedback scheme for a large number of transmit antennas.

We can model temporally correlated channels as a first-order Gauss-Markov process as [27]

$$\mathbf{h}[k] = \eta\mathbf{h}[k-1] + \sqrt{1-\eta^2}\mathbf{g}[k] \tag{12}$$

where η $(0 \leq \eta \leq 1)$ is a temporal correlation coefficient which represents the correlation between elements $h_m[k-1]$ and $h_m[k]$ where $h_m[k]$ is the m^{th} entry of $\mathbf{h}_m[k]$. Using Jakes' model [28], $\eta = J_0(2\pi f_D T)$, where $J_0(\cdot)$ is the 0th order Bessel function of the first kind, f_D denotes the maximum Doppler frequency and τ denotes the channel instantiation interval. $\mathbf{g}[k]$ is the evolution process having i.i.d. entries distributed with $\mathcal{CN}(0, 1)$. We assume that the initial state $\mathbf{h}[0]$ is independent of $\mathbf{g}[k]$ for all k.

As shown in Fig. 1, the function $Q(\cdot)$ maps the channel vector $\mathbf{h}[k]$ to the binary vector $\mathbf{b}[k]$ and $R(\cdot)$ reconstructs $\mathbf{h}[k]$ to $\mathbf{f}[k]$. Therefore, we denote $R(Q(\cdot))$ as the NTCQ scheme explained in Section III and $\mathbf{f}[k-1] = R(Q(\bar{\mathbf{h}}[k-1]))$ as the beamforming vector at time $k-1$. Instead of quantizing $\bar{\mathbf{h}}[k]$ directly at time k, the receiver first projects $\bar{\mathbf{h}}[k]$ to the null space of $\mathbf{f}[k-1]$ as

$$\mathbf{f}_{\mathrm{null}}[k] = \left(\mathbf{I}_{M(k)} - \mathbf{f}[k-1]\mathbf{f}[k-1]^H\right)\bar{\mathbf{h}}[k]. \tag{13}$$

The receiver quantizes $\mathbf{f}_{\mathrm{null}}[k]$ as $\hat{\mathbf{f}}_{\mathrm{null}}[k] = R(Q(\mathbf{f}_{\mathrm{null}}[k]))$ and constructs candidate beamforming vectors $\mathbf{f}_{\bar{\alpha},\bar{\theta}}$ using weights $\bar{\alpha} \in \{\bar{\alpha}_1, \ldots, \bar{\alpha}_{K_{\bar{\alpha}}}\} = \bar{\mathbb{A}}$ and $\bar{\theta} \in \{\bar{\theta}_1, \ldots, \bar{\theta}_{K_{\bar{\theta}}}\} = \bar{\Theta}$ as

$$\mathbf{f}_{\bar{\alpha},\bar{\theta}} = \frac{\eta\mathbf{f}[k-1] + \bar{\alpha}e^{j\bar{\theta}}\sqrt{1-\eta^2}\hat{\mathbf{f}}_{\mathrm{null}}[k]}{\|\eta\mathbf{f}[k-1] + \bar{\alpha}e^{j\bar{\theta}}\sqrt{1-\eta^2}\hat{\mathbf{f}}_{\mathrm{null}}[k]\|_2}. \tag{14}$$

The receiver selects the pair of the best combiners $(\bar{\alpha}_{\mathrm{best}}, \bar{\theta}_{\mathrm{best}})$ by optimizing

$$\max_{\bar{\alpha}\in\bar{\mathbb{A}}} \max_{\bar{\theta}\in\bar{\Theta}} |\bar{\mathbf{h}}^H \mathbf{f}_{\bar{\alpha},\bar{\theta}}|^2, \tag{15}$$

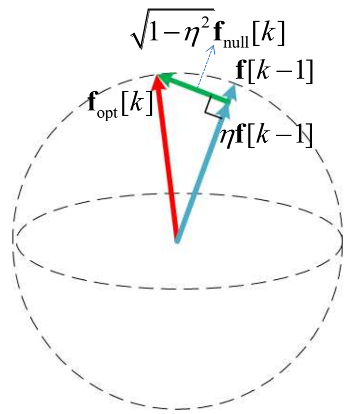

Fig. 2: Conceptual explanation of the proposed differential scheme with NTCQ

and the final beamforming vector $\mathbf{f}[k]$ becomes

$$\mathbf{f}[k] = \mathbf{f}_{\bar{\alpha}_{\text{best}},\bar{\theta}_{\text{best}}}. \tag{16}$$

The conceptual explanation of this differential scheme is shown in Fig. 2. We need two sets of combiners $\bar{\mathbb{A}}$ and $\bar{\Theta}$ to construct candidate beamformaing vectors as in (14). It is obvious to have $\bar{\theta} \in [0, 2\pi)$. We make the following proposition to get the range of $\bar{\alpha}$.

Proposition 1. *For temporally correlated channels, $\bar{\alpha}$ in (14) can be bounded as*

$$\frac{1-\eta}{\sqrt{1-\eta^2}} \leq \bar{\alpha} \leq \frac{1+\eta}{\sqrt{1-\eta^2}} \tag{17}$$

when $\eta \to 1$.

Proof: First, expand the norm square of the denominator of (14) as

$$\|\eta\mathbf{f}[k-1] + \bar{\alpha}e^{j\bar{\theta}}\sqrt{1-\eta^2}\hat{\mathbf{f}}_{\text{null}}[k]\|_2^2 = \eta^2 + \bar{\alpha}^2(1-\eta^2) + 2\bar{\alpha}\eta\sqrt{1-\eta^2}\text{Re}\left\{e^{j\bar{\theta}}\mathbf{f}^H[k-1]\hat{\mathbf{f}}_{\text{null}}[k]\right\}. \tag{18}$$

Because $||\mathbf{f}[k-1]||_2^2 = 1$ and $||\mathbf{f}_{\text{null}}[k]||_2^2 = 1$, we have

$$-1 \leq \text{Re}\left\{e^{j\bar{\theta}}\mathbf{f}^H[k-1]\hat{\mathbf{f}}_{\text{null}}[k]\right\} \leq 1,$$

$$\left(\eta - \bar{\alpha}\sqrt{1-\eta^2}\right)^2 \leq (18) \leq \left(\eta + \bar{\alpha}\sqrt{1-\eta^2}\right)^2.$$

With a good quantizer $R(Q(\cdot))$, we can approximate $\mathbf{f}^H[k-1]\hat{\mathbf{f}}_{\text{null}}[k] \approx 0$. Moreover, we can approximate $\eta \approx 1$ because of the assumption of a slowly time-varying channel for differential codebooks. Then, we have (18) $= 1$ and the range of $\bar{\alpha}$ becomes (17). ■

The range in (17) can be further optimized, however, optimizing the range of $\bar{\alpha}$ is out of the scope of this paper. We set $\frac{1-\eta}{\sqrt{1-\eta^2}} \leq \bar{\alpha} \leq \frac{1+\eta}{3\sqrt{1-\eta^2}}$ for simulation purpose in Section V.

After the receiver selects the beamforming vector $\mathbf{f}[k]$ as in (16), it has to feed back $\hat{\mathbf{f}}_{\text{null}}[k]$, $\bar{\alpha}_{\text{best}}$, and $\bar{\theta}_{\text{best}}$ to the transmitter. Because of the information of $\bar{\alpha}_{\text{best}}$, and $\bar{\theta}_{\text{best}}$, feedback overhead is increased in the proposed differential scheme. However, the additional feedback overhead of $\bar{\alpha}_{\text{best}}$ and $\bar{\theta}_{\text{best}}$ can be marginal compared to that of $\hat{\mathbf{f}}_{\text{null}}[k]$. Our simulations show that it is enough to have 1bit for $\bar{\alpha}_{\text{best}}$ and 3bits for $\bar{\theta}_{\text{best}}$ to achieve near-optimal performance. Once the transmitter receives the feedback information from the receiver, it reconstructs $\mathbf{f}[k]$ as in (14).

B. Adaptive NTCQ Scheme

In massive MIMO systems, the transmit antennas have to be deployed in a potentially limited area resulting in channels that may be spatially correlated. The spatially correlated channel is often modeled as

$$\mathbf{h}[k] = \mathbf{R}^{\frac{1}{2}}\mathbf{h}_w[k] \tag{19}$$

where $\mathbf{R}$ is an invertible, long-term spatial correlation matrix and $\mathbf{h}_w[k]$ is a uncorrelated MISO channel vector with i.i.d. complex Gaussian entries.

In [17]–[20], codebook adaptation methods for spatially correlated MISO channels have been proposed. These works roughly quantize only the local area of the dominant eigenvector of $\mathbf{R}$ by rotating and normalizing codewords in a fixed codebook with respect to $\mathbf{R}$. It has been shown that this adaptation method can reduce the channel quantization error significantly with the same number of codewords. However, candidate codewords for channel quantization are not easily fixed in NTCQ so it is not possible to apply the adaptation method directly. Therefore, we come up with the following method for NTCQ to mimic the adaptation methods in [17]–[19] for spatially correlated MISO channels.

We assume that the transmitter and the receiver both know $\mathbf{R}$ perfectly because it is a long-term channel statistic. The receiver first decorrelates $\mathbf{h}[k]$ with $\mathbf{R}^{-\frac{1}{2}}$ and gets $\mathbf{h}_w[k]$ as

$$\mathbf{h}_w[k] = \mathbf{R}^{-\frac{1}{2}}\mathbf{h}[k]. \tag{20}$$

Then the receiver quantizes $\mathbf{h}_w[k]$ as $\hat{\mathbf{h}}_w[k] = R(Q(\mathbf{h}_w[k]))$ and feeds back $\hat{\mathbf{h}}_w[k]$ to the transmitter. The transmitter reconstructs the beamforming vector $\mathbf{f}[k]$ as

$$\mathbf{f}[k] = \frac{\mathbf{R}^{\frac{1}{2}}\hat{\mathbf{h}}_w[k]}{||\mathbf{R}^{\frac{1}{2}}\hat{\mathbf{h}}_w[k]||_2}. \tag{21}$$

The procedures in (20) and (21) decouple the correlation and the quantization part effectively and gives the same result as the adaptation method for fixed codewords.

V. Simulations

In this section, we verify the performance of the advanced NTCQ schemes in temporally and spatially correlated channels by Monte Carlo simulation. The total number of feedback bits for each scheme is shown in Table I. Each number in the table is the required feedback bits for $M(k) = 5$ / 10 / 20 transmit antennas. Note that 1, 2, and 3 bits/entry correpond to NTCQ

TABLE I: Summary of feedback overhead

	1bit/entry	2bits/entry	3bits/entry
Orig. NTCQ	7 / 12 / 22	13 / 23 / 43	18 / 33 / 63
Diff. NTCQ	11 / 16 / 26	17 / 27 / 47	22 / 37 / 67
Adap. NTCQ	7 / 12 / 22	13 / 23 / 43	18 / 33 / 63

with QPSK, 8PSK, and 16QAM constellation, respectively. We define the average normalized received SNR $\bar{\Gamma}_{\text{avg}}$ as

$$\bar{\Gamma}_{\text{avg}} = \frac{\rho}{\sigma^2} E[\Gamma(\mathbf{f}[k])] = E[|\mathbf{h}[k]^H \mathbf{f}[k]|^2] \qquad (22)$$

and use $\bar{\Gamma}_{\text{avg}}$ as a performance metric. Note that the expectation in (22) is over $\mathbf{h}$.

A. Temporally Correlated Channels

We assume a $2.5GHz$ carrier frequency and $\tau = 5ms$ channel instantiation interval. We uniformly quantize the combiners $\bar{\theta}$ and $\bar{\alpha}$ in (14) with 3bits and 1bit, respectively.

We plot the performance of differential NTCQ schemes with $v = 3km/h$ ($\eta = 0.9881$) and $v = 10km/h$ ($\eta = 0.8721$) in Fig. 3 and 4 to see the effect of channel variation. Even with 1 bit/entry quantization, the differential NTCQ schemes give almost the same performance as unquantized beamforming in a $v = 3km/h$ environment regardless of the number of transmit antennas and still achieve a large fraction of unquantized beamforming gain in a $v = 10km/h$ environment.

B. Spatially Correlated Channels

To simulate spatially correlated channels, we adopt the Kronecker model for the correlation matrix $\mathbf{R}$ which is written as

$$\mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H \qquad (23)$$

where $\mathbf{U}$ and $\mathbf{\Sigma}$ are $M(k) \times M(k)$ eigenvector and diagonal eigenvalue matrices, respectively. We assume $\mathbf{\Sigma}$ has a structure

$$\mathbf{\Sigma} = \text{diag}\left\{\lambda_1, \frac{M(k)-\lambda_1}{M(k)-1}, \cdots, \frac{M(k)-\lambda_1}{M(k)-1}\right\} \qquad (24)$$

where $1 \leq \lambda_1 < M(k)$ is the maximum eigenvalue of $\mathbf{R}$ to model varying amounts of spatial correlation. If λ_1 is large (small), then channels are highly (loosely) correlated in spatial domain.

In Figs. 5, and 6, we plot $\bar{\Gamma}_{\text{avg}}$ with different values of λ_1 for $M(k) = 10$ and 20 cases. In both values of $M(k)$, the performance of adaptive NTCQ schemes becomes closer to that of unquantized beamforming with the same feedback overhead as the original NTCQ as λ_1 increases.

Simulation results in this section show that utilizing statistical information for temporally/spatially correlated channels is crucial to increase the performance or decrease the feedback overhead.

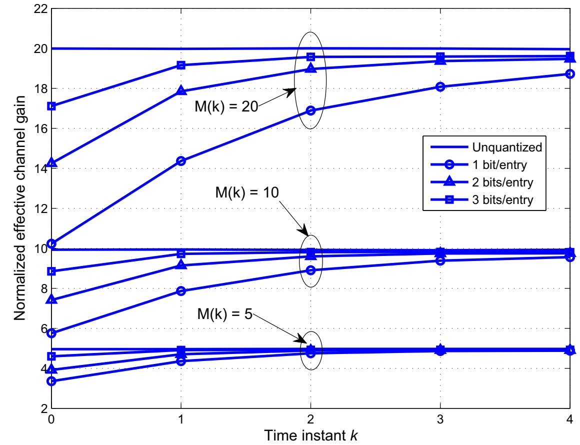

Fig. 3: Plot of $\bar{\Gamma}_{\text{avg}}$ versus k in temporally correlated channels for $v = 3km/h$ ($\eta = 0.9881$).

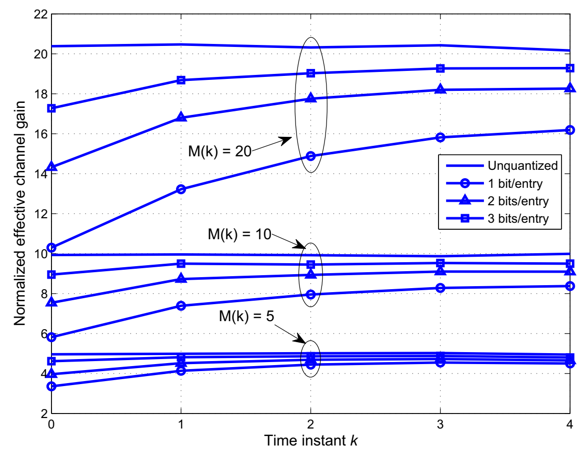

Fig. 4: Plot of $\bar{\Gamma}_{\text{avg}}$ versus k in temporally correlated channels for $v = 10km/h$ ($\eta = 0.8721$).

VI. Conclusions

The results in this paper demonstrate that advanced noncoherent trellis-coded quantization strategies that account for temporal and spatial correlations in the channel provide an effective approach for obtaining channel state information in massive MIMO systems with moderate feedback overhead. The differential and spatially adaptive schemes proposed here each improve upon the original NTCQ scheme, and an interesting topic for future research is to investigate additional performance improvements from combining these schemes.

References

[1] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput*, IEEE Std., 2009.

[2] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation*, 3GPP TS 36.211 v11.0.0 Std., 2012.

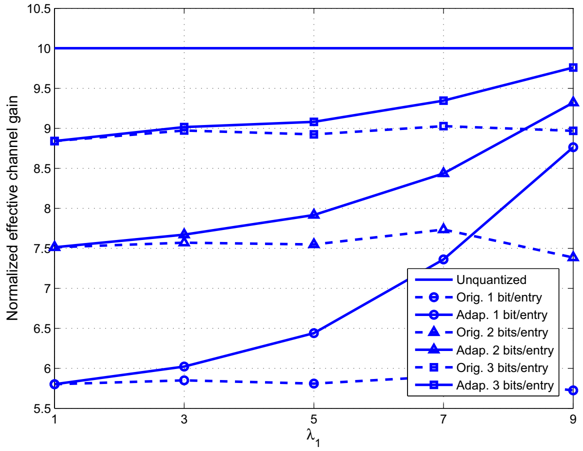

Fig. 5: Plot of $\bar{\Gamma}_{\mathrm{avg}}$ versus λ_1 in spatially correlated channels, $M(k) = 10$.

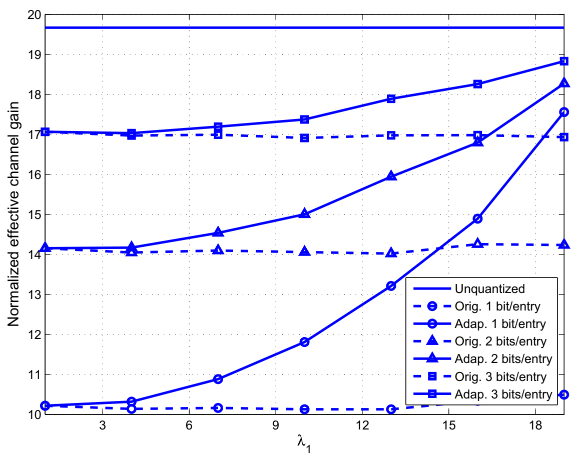

Fig. 6: Plot of $\bar{\Gamma}_{\mathrm{avg}}$ versus λ_1 in spatially correlated channels, $M(k) = 20$.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] ——, "How much training is required for multiuser MIMO?" *Proceedings of IEEE Asilomar Conference on Signals, Systems, and Computers*, Oct. 2006.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Uplink power efficiency of multiuser MIMO with very large antenna arrays," *Proceedings of Allerton Conference on Communication, Control, and Computing*, Sep. 2011.

[6] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO: how many antennas do we need?" *Proceedings of Allerton Conference on Communication, Control, and Computing*, Sep. 2011.

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

[8] J. Choi, Z. Chance, D. J. Love, and U. Madhow, "Noncoherent trellis coded quantization for massive MIMO limited feedback beamforming," *UCSD Information Theory and Applications Workshop*, Feb. 2013.

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

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

[11] C. Simon and G. Leus, "Feedback quantization for linear precoded spatial multiplexing," *EURASIP Journal on Advances in Signal Processing*, article ID 683030, 13 pages, 2008.

[12] R. W. Heath Jr., T. Wu, and A. C. K. Soong, "Progressive refinement of beamforming vectors for high-resolution limited feedback," *EURASIP Journal on Advances in Signal Processing*, article ID 463823, 13 pages, 2009.

[13] D. Sacristan and A. Pascual-Iserte, "Differential feedback of MIMO channel gram matrices based on geodesic curves," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3714–3727, Dec. 2010.

[14] T. Kim, D. J. Love, and B. Clerckx, "MIMO system with limited rate differential feedback in slow varying channel," *IEEE Transactions on Communications*, vol. 59, no. 4, pp. 1175–1180, Apr. 2010.

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

[16] J. Choi, B. Clerckx, and D. J. Love, "Differential codebook for general rotated dual-polarized MISO channels," *Proceedings of IEEE Global Telecommunications Conference*, Dec. 2012.

[17] D. J. Love and R. W. Heath Jr., "Grassmannian beamforming on correlated MIMO channels," *Proceedings of IEEE Global Telecommunications Conference*, Dec. 2004.

[18] ——, "Limited feedback diversity techniques for correlated channels," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 718–722, Mar. 2006.

[19] P. Xia and G. B. Giannakis, "Design and analysis of transmit-beamforming based on limited-rate feedback," *IEEE Transactions on Signal Processing*, vol. 54, no. 5, pp. 1853–1863, Mar. 2006.

[20] V. Raghavan, R. W. Heath Jr., and A. M. Sayeed, "Systematic codebook designs for quantized beamforming in correlated MIMO channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 7, pp. 1298–1310, Sep. 2007.

[21] C. K. Au-Yeung, D. J. Love, and S. Sanayei, "Trellis coded line packing: large dimensional beamforming vector quantization and feedback transmission," *IEEE Transactions on Wireless Communications*, vol. 10, no. 6, pp. 1844–1853, Jun. 2011.

[22] C. K. Au-Yeung and S. Sanayei, "Enhanced trellis based vector quantization for coordinated beamforming," *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, Mar. 2010.

[23] M. Xu, D. Guo, and M. K. Honig, "MIMO precoding with limited rate feedback: simple quantizers work well," *Proceedings of IEEE Global Telecommunications Conference*, Dec. 2009.

[24] D. J. Ryan, I. V. L. Clarkson, I. B. Collings, D. Guo, and M. L. Honig, "QAM and PSK codebooks for limited feedback MIMO beamforming," *IEEE Transactions on Communications*, vol. 57, no. 4, pp. 1184–1196, Apr. 2009.

[25] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Transactions on Information Theory*, pp. 55–67, Jan. 1982.

[26] D. Warrier and U. Madhow, "Spectrally efficient noncoherent communication," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 652–668, Mar. 2002.

[27] R. H. Etkin and D. N. C. Tse, "Degree of freedom in some underspread MIMO fading channel," *IEEE Transactions on Information Theory*, vol. 52, pp. 1576–1608, Apr. 2006.

[28] J. G. Proakis, *Digital communication*, 4th ed. New York: McGraw-Hill, 2000.