Channel Vector Quantization for Multiuser MIMO Systems Aiming at Maximum Sum Rate

Guido Dietl and Olivier Labrèche
DOCOMO Communications Laboratories Europe GmbH
Landsberger Str. 312, 80687 Munich, Germany
Email: dietl@docomolab-euro.com

Wolfgang Utschick
Associate Institute for Signal Processing
Technische Universität München
80333 Munich, Germany

Abstract—For downlink transmission in a multiuser Multiple-Input Multiple-Output (MIMO) communication system, quantized Channel State Information (CSI) is fed back to the base station in an uplink channel of finite rate. The quantized CSI is obtained via Channel Vector Quantization (CVQ) of the so-called composite channel vector, i.e., the product of the channel matrix and an estimation of the receive filter, which cannot be computed exactly at the stage of quantization because of its dependency on the finally chosen precoder. Here, the state-of-the-art approach estimates the receive filter and quantizes the composite channel vector such that its Euclidean distance to the estimated composite channel vector is minimized.

In this paper, we propose an alternative CVQ method which determines the estimated receive filter vector and the quantized composite channel vector such that the resulting Signal-to-Interference-and-Noise Ratio (SINR), or an approximation thereof, is maximized. Since the SINR is related to the individual user rates, and therefore related to the sum rate of the system, the presented solution aims at maximizing the system sum rate.

Simulation results of a multiuser MIMO system with linear zero-forcing precoding show that the proposed schemes achieve significant performance improvements compared to the state-of-the-art method, especially in the low signal-to-noise ratio region.

I. INTRODUCTION

We consider the downlink of a multiuser Multiple-Input Multiple-Output (MIMO) system, also known as the MIMO broadcast channel, where the base station transmits data symbols over multiple antennas to users with multiple receive antennas. Recently, it has been shown in [1] that nonlinear dirty paper coding achieves the Sato bound [2], i.e., the upper bound on the sum capacity of the MIMO broadcast channel, revealing its high capacity potential. It followed the publication of many nonlinear precoding algorithms (e.g., [3], [4]) which are designed to reach or be very close to the maximum sum capacity if the transmitter has perfect CSI.

In this paper, we focus on systems where the transmitter has only partial CSI via a limited feedback channel from the mobile stations, i.e., only a finite number of bits can be used for the information fed back to the base station. The consideration of such systems assuming nonlinear transmit processing can be found in, e.g., [5], [6]. However, we are especially interested in linear precoders because of their implementation advantages over nonlinear schemes like being in general computationally more efficient, having smaller processing delays by avoiding successive encoding, and inducing less requirements on hardware like the dynamic range of amplifiers or analog-to-digital converters. Compared to the investigation of linear precoders in [7], [8] (see also references therein) for a single-user system, or in [9] (see also references therein) for a multiuser system, we consider a multiuser system with user scheduling in order to fully exploit multiuser diversity.

The feedback approach is based on Channel Vector Quantization (CVQ) using a finite channel codebook as proposed in [10]–[12]. Based on this codebook, each user quantizes a product of his channel matrix and an estimation of its receive filter, in the following denoted as the composite channel vector, and feeds back the corresponding codebook index together with an approximate Signal-to-Interference-and-Noise Ratio (SINR) value. Note that users need to estimate their receivers because the finally chosen receive filters depend on the precoder at the base station which is determined after quantization. In fact, the base station uses the quantized composite channel vectors to compute the precoder based on the Zero-Forcing (ZF) criterion, and uses the available SINR values to schedule the users by maximizing the sum rate.

Usually, CVQ is based on choosing the codebook entry with minimum Euclidean distance to the composite channel vector. However, minimizing the Euclidean distance is not necessarily related to the final goal of designing a communications system, i.e., maximizing the sum rate. Therefore, we propose to estimate the receive filter and quantize the corresponding composite channel vector by maximizing the approximate SINR which is directly related to the achievable rate of the corresponding user.

Before reviewing the state-of-the-art CVQ methods and deriving the proposed CVQ approaches in Section III, we introduce the model of the multiuser MIMO system in the next section. Finally, we investigate the performance of the proposed schemes when applied to a multiuser MIMO system with linear ZF precoding in Section IV.

Throughout the paper, vectors and matrices are denoted by lower and upper case bold letters, respectively. The matrix $I_n$ is the $n \times n$ identity matrix, $e_v$ its $v$th column whose dimension is defined implicitly by the context, and $0_n$ the $n$-dimensional zero vector. The operation $(\cdot)^*$ denotes conjugate, $(\cdot)^T$ transpose, $(\cdot)^H$ Hermitian, i.e., conjugate transpose, $\| \cdot \|_2$ the Euclidean norm, and $| \cdot |$ the cardinality of a set or the absolute value of a scalar, respectively. $N_c(m, C)$ denotes a multivariate proper complex normal distribution with mean...
vector \( m \) and covariance matrix \( C \).

II. MULTIUSER MIMO SYSTEM MODEL

A. Multiuser MIMO Downlink Channel

In the downlink of a multiuser MIMO system as depicted in Fig. 1, the scheduled and precoded symbols, summarized in the vector \( x \sim \mathcal{N}(0_M, C_x) \), are transmitted via the \( M \) antennas of the base station to the \( K \) mobile stations (users), each equipped with \( N \) receive antennas. Note that the transmit power at the base station is \( P_{tx} = \text{tr}(C_x) \). With the \( k \)-th user’s channel matrix \( H_k \in \mathbb{C}^{N \times M} \), \( k \in \{1, \ldots, K\} \), and the noise vector \( n_k \sim \mathcal{N}(0_N, I_N) \) with independent and normal distributed elements of variance one, the perturbed receive vector of the \( k \)-th user can be written as

\[
y_k = H_k x + n_k \in \mathbb{C}^N. \tag{1}
\]

B. Linear Precoder and Receivers

In the following, we assume that maximal one data stream is assigned to each user. Consequently, the number of scheduled users is equal to the number of scheduled data streams, i.e., \( |K| = D \). Further, let us assume that the scheduler has already selected the users for transmission which are summarized in the set \( K \subseteq \{1, \ldots, K\} \). Then, the symbols \( s_k \sim \mathcal{N}(0, P_{tx}/D) \) of all selected users \( k \in K \) are precoded using the precoders \( p_k \in \mathbb{C}^M \), and summed up to get the transmit vector

\[
x = \sum_{k \in K} p_k s_k = P s. \tag{2}
\]

Here, \( P \in \mathbb{C}^{M \times D} \) is a matrix where the precoders \( p_k, k \in K \), are stacked in a row with ascending \( k \), in the following denoted as \( P = \text{rstack}(p_k)_{k \in K} \), and \( s \in \mathbb{C}^D \) is a vector where the symbols \( s_k, k \in K \), are stacked in a column with ascending \( k \), in the following denoted as \( s = \text{cstack}(s_k)_{k \in K} \).

Next, we describe the receivers at the mobile stations. We assume that each user is applying a linear filter \( w_k \in \mathbb{C}^N \) to the receive vector \( y_k \) to get the estimate

\[
\hat{s}_k = w_k^H y_k \in \mathbb{C}, \tag{3}
\]

of the \( k \)-th user’s symbol \( s_k \). In particular, we consider the linear Minimum Mean Square Error (MMSE) filter obtained via the minimization of the mean-square error between \( s_k \) and \( \hat{s}_k = w_k^H y_k \), whose solution computes as (e.g., [13])

\[
w_k = \left( H_k^H P^H P^H + \frac{D}{P_{tx}} I_N \right)^{-1} H_k^H p_k. \tag{4}
\]

Note that we assume throughout this paper that each receiver has access to the perfect Channel State Information (CSI) of its own channel \( H_k \). However, it has no CSI about the channels of the other users due to the non-cooperative nature of the multiuser MIMO downlink channel.

C. Sum Rate Performance Measure

A typical measure for the downlink transmission performance of a multiuser MIMO system is the sum rate over all users. With the assumptions made in Sections II-A and II-B, the Signal-to-Interference-and-Noise Ratio (SINR) at the output of the receive filter \( w_k \) of the \( k \)-th user can be written as

\[
\gamma_k = \frac{|w_k^H H_k p_k|^2}{\|w_k\|_2^2 P_{tx} + \sum_{i \in K} \sum_{j \neq k} |w_i^H H_k p_j|^2}, \tag{5}
\]

and the sum rate computes as

\[
R_{\text{sum}} = \sum_{k \in K} \log_2 (1 + \gamma_k). \tag{6}
\]

Remember that the variance of each user’s symbol is set to \( P_{tx}/D \).

III. CHANNEL VECTOR QUANTIZATION (CVQ)

In order to compute the precoder and schedule the users for transmission, the base station requires information about the channel matrices \( H_k \) for all \( k \in \{1, \ldots, K\} \). This CSI is fed back from the terminals to the base station. Precisely speaking, each user quantizes its channel based on a channel codebook and feeds back the corresponding codebook index together with an SINR value which includes a rough estimate of the interference caused by the quantization error [9]–[11], [14]. The base station is then computing a precoder, e.g., of type Zero-Forcing (ZF), based on the quantized CSI and allocates resources using the available SINR information. Again, since we are especially interested in low feedback schemes, we restrict the maximum number of transmitted data symbols per user to one.

Assume for the first that the precoder \( P \) is known at the mobile receivers such that the MMSE filters \( w_k \in \mathbb{C}^N \) can be computed according to Eq. (4). In order to compute the feedback information, each user \( k \) quantizes the composite channel vector \( g_k = H_k^T w_k \in \mathbb{C}^M \), being a combination of the linear MMSE filter and the physical channel matrix, by applying Channel Vector Quantization (CVQ). With the channel codebook \( \mathcal{C} = \{u_1, \ldots, u_{2^B}\} \) where \( B \) denotes the number of necessary bits for indexing the \( 2^B \) normalized codebook vectors \( u_q \in \mathbb{C}^M, q \in \{1, \ldots, 2^B\} \).

A. Minimum Euclidean Distance Based CVQ

The quantized composite channel vector \( \hat{g}_k \in \mathbb{C}^M \) is the codebook entry which has the minimum Euclidean distance to the normalized composite channel vector (minimum quantization angle/error), i.e., for \( k \in K \):

\[
\hat{g}_k = u_\ell, \quad \ell = \arg\max_{q \in \{1, \ldots, 2^B\}} \frac{|u_q^H g_k|}{\|g_k\|_2}. \tag{7}
\]
Here, $\ell$ denotes the codebook index which is fed back to the base station using $B$ bits. Note that the codebook index provides only the direction of the channel vector whereas the magnitude is included in the SINR value which is also part of the feedback information. As derived in [10], [14], a scaled version of the SINR at the $k$th mobile receiver is approximated via

$$
\gamma_k' = \frac{P_{\text{tx}}}{1 + \frac{P_{\text{tx}}}{2} \|g_k\|^2 \sin^2 \theta_k} \cos \theta_k = \frac{|g_k^H g_k|}{\|g_k\|^2},
$$

where $\theta_k \in [0, \pi]$ denotes the angle between the normalized composite channel vector and the quantized version thereof (quantization angle), and where, without loss of generality, we set $\|w_k\|_2 = 1$. Again, note that the scaled SINR value in Eq. (8) includes an approximation of the interference caused by the Zero-Forcing (ZF) approach due to the quantization error (cf. Section IV).

It remains to explain how the user can approximate the composite channel vector $g_k$ without knowing the precoder $P$ at the base station. We follow the ideas of [9] where the composite channel is chosen as a linear combination of the rows of $H_k$ such that the quantization error is minimized (cf. Figure 2). With the QR factorization of $H_k^T = Q_k R_k$ where $Q_k \in \mathbb{C}^{M \times N}$ is a matrix with orthonormal columns and $R_k \in \mathbb{C}^{N \times N}$ is upper triangular, the CVQ of the composite channel vector is obtained by choosing the codebook entry in $C$ which is closest to the row space of $H_k$ with respect to the Euclidean Distance (ED):

$$
g_k^\text{ED} \triangleq \arg\max_{u_v \in C} \|Q_k^H u_v\|^2_2.
$$

Then, the approximation of the composite channel vector is obtained by projecting $g_k^\text{ED}$ back into the row space of $H_k$

$$
g_k = \frac{Q_k Q_k^H g_k^\text{ED}}{\|Q_k Q_k^H g_k^\text{ED}\|^2_2},
$$

and applying the scaling

$$
g_k^\text{ED} \triangleq \frac{g_k}{\|g_k H_k^\dagger\|^2_2},
$$

where $H_k^\dagger = H_k^H (H_k H_k^H)^{-1}$ is the right-hand side pseudo-inverse of the matrix $H_k$. The latter scaling is necessary to ensure that $\|w_k^\text{ED}\|_2 = \|g_k^\text{ED, T} H_k^\dagger\|^2_2 = 1$. Replacing $\hat{g}_k$ and $\hat{g}_k$ in Eq. (8) by $g_k^\text{ED}$ of Eq. (9) and $\hat{g}_k^\text{ED}$ of Eq. (11), respectively, yields the scaled SINR value $\gamma_k'$ which is the additional feedback information to the codebook index.

B. Maximum SINR Based CVQ

In this paper, we propose an alternative CVQ method which is based on the maximization of the SINR expression. The original CVQ approach from [9] is only concerned with minimizing the quantization error represented by the quantization angle $\theta_k$. However, the ultimate objective in our system should be to targeted the highest possible sum rate. Due to Eqs. (5) and (6), this is achieved by maximizing the SINR, which can be approximated via expressions such as Eq. (8).

Due to the above reasons, we argue that an approach based on the maximization of an SINR expression may prove to yield superior performance in a sum rate sense than simply minimizing the quantization angle. Our method is solely concerned in finding the best effective channel $\hat{g}_k$ and its quantized version $\hat{g}_k$. In that sense, it provides an alternative to Section III-A.

Let us have a look again at the approximate expression for the scaled SINR in Eq. (8). Note that this SINR expression will approach $\cot^2 \theta_k$ as $P_{\text{tx}}$ goes to infinity. Therefore, at large SNR, the SINR is almost fully determined by the quantization error, which justifies the minimization of $\theta_k$ in such conditions. However, the situation might be different at lower SNRs.

Thus, our proposed approach is to maximize Eq. (8) over all possible codebook entries $g_k$ and receiver weights $w_k$ of unit norm, i.e.,

$$
\left(\mathbf{w}_{k}^{\text{SINR}}, \mathbf{g}_{k}^{\text{SINR}}\right) = \arg\max_{(w_k, g_k) \in \{\mathbb{C}^N : \|w_k\|_2 = 1\} \times C} \gamma_k'(w_k, g_k),
$$

yielding us an optimal $g_k^{\text{SINR}}$ and $w_k^{\text{SINR}}$. The optimal composite channel $g_k^{\text{SINR}}$ is then computed as $g_k^{\text{SINR}} = H_k^T w_k^{\text{SINR}}$. The resulting quantities are the ones finally used for the SINR computations according to Eq. (8) by replacing $\hat{g}_k$ and $g_k$ with $g_k^{\text{SINR}}$ and $g_k^{\text{SINR}}$, respectively.

In order to perform the above maximization, let us rewrite Eq. (8) by substituting $\|g_k^\dagger\|^2 \cos^2 \theta_k = |g_k^H g_k|^2$ as $g_k^{\text{SINR}} = H_k^T w_k^{\text{SINR}}$. (unit norm constraint of receive filters):

$$
\gamma_k'(w_k, g_k^{\text{SINR}}) = \frac{(w_k^H H_k^\dagger g_k^{\text{SINR}} H_k^\dagger)^T w_k}{w_k^H (I + \frac{P_{\text{tx}}}{2} H_k^\dagger (I - g_k^H g_k^H) H_k^\dagger) w_k} = \frac{(w_k^H A g_k)^T w_k}{w_k^H B(g_k) w_k}.
$$

It is well known that expressions in the form of (13) are maximized by setting $w_k$ to the eigenvector $v_i$ corresponding to the largest eigenvalue $\mu_i$ solving the generalized eigenvalue problem $A v_i = \mu_i B v_i$ (e.g., [15]). Moreover, if $B$ is invertible, the eigenvalues and eigenvectors are the same as

978-1-4244-4148-8/09/$25.00 ©2009
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE "GLOBECOM" 2009 proceedings.
for the regular eigenvalue decomposition of $B^{-1}A$.

Note that this maximization finds the best $w_k$ given a specific codebook entry $g_k$. The optimal $g_k^{\text{SINR}}$ is the one yielding the largest SINR value over all codebook entries, i.e.,

$$g_k^{\text{SINR}} = \underset{g_k \in \mathcal{C}}{\operatorname{argmax}} \max_{w_k \in \{\mathbb{C}^N: \|w_k\|_2 = 1\}} \frac{w_k^H \Gamma g_k w_k}{w_k^H B g_k w_k}, \quad (14)$$

and $w_k^{\text{SINR}}$ is the corresponding optimal weight vector for this $g_k^{\text{SINR}}$. A drawback of this method applied directly is the computational complexity, since the maximization over $w_k$ needs to be performed for all entries of the channel codebook, and thus, requires $2^B$ generalized eigenvalue decompositions per sub-carrier where the optimization is performed.

**C. Pseudo-Maximum SINR Based CVQ**

To overcome the high computational burden of the maximum SINR solution, we propose a pseudo-maximization algorithm as an alternative to the exact maximization in Eq. (14). We note that $\gamma_k^*$ is only a function of the Channel Magnitude (CM) $\|g_k\|_2^2$ and the quantization error $\delta_k$ for a given $P_{\text{tx}}$. The only way to increase $\gamma_k^*$ is by increasing $\|g_k\|_2$ or decreasing $\delta_k$. The critical assumption we will make here is that $\gamma_k^*$ is close to its maximum when either $\|g_k\|_2$ is maximized or $\delta_k$ is minimized. Therefore, we evaluate $\gamma_k^*$ at two specific points, denoted as $(w_k^{\text{CM}}, g_k^{\text{CM}})$ and $(w_k^{\text{ED}}, g_k^{\text{ED}})$, where we define

$$w_k^{\text{CM}} = \arg \max_{w_k \in \{\mathbb{C}^N: \|w_k\|_2 = 1\}} \|g_k^{\text{CM}}\|_2^2$$

$$= \arg \max_{w_k \in \{\mathbb{C}^N: \|w_k\|_2 = 1\}} w_k^H H_k^* H_k^T w_k,$$

$$g_k^{\text{CM}} = \arg \max_{g_k \in \mathcal{C}} \|g_k^{\text{CM}}\|_2^2$$

$$= \arg \max_{g_k \in \mathcal{C}} g_k^H H_k^* H_k^T w_k^{\text{CM}},$$

and the point $(w_k^{\text{ED}}, g_k^{\text{ED}})$ is the minimum Euclidean distance solution that we already presented in Section III-A. Note that $w_k^{\text{CM}}$ is in the direction of the eigenvector corresponding to the largest eigenvalue of $H_k^* H_k^T$. The increase in computational complexity of the pseudo-maximization solution with respect to the minimum Euclidean distance method therefore includes one (regular) eigenvalue decomposition and one search for the closest quantization vector $g_k^{\text{CM}}$.

We finally compute the Pseudo-Maximum (PM) solution to the SINR maximization problem as

$$\left( w_k^{\text{PM}}, g_k^{\text{PM}} \right) = \underset{(w_k, g_k) \in \{(w_k^{\text{CM}}, g_k^{\text{CM}}), (w_k^{\text{ED}}, g_k^{\text{ED}})\}}{\arg \max} \gamma_k^*(w_k, g_k). \quad (16)$$

**IV. APPLICATION TO ZERO-FORCING (ZF) PRECODING**

**A. ZF Precoding**

With the quantized composite channel matrix $\hat{G}_K = \text{cstack}(\hat{g}_k^T)_{k \in \mathbb{K}} \in \mathbb{C}^{D \times M}$, the ZF precoder at the base station computes as (e.g., [13])

$$P_{K} = \hat{P} K A_{K}^{1/2}, \quad \hat{P}_K = \hat{G}_K^H \left( \hat{G}_K \hat{G}_K^H \right)^{-1}, \quad (17)$$

with the diagonal matrix $A_K \in \mathbb{C}^{D \times D}$ representing power loading. For the simulations in Section IV-C, we assume equal power loading where

$$A_K = \text{diag} \left( \frac{P_{\text{tx}}}{D \|P_{K}^* e_k\|^2_{/k=1}} \right). \quad (18)$$

Recall that the base station assigns maximal one data stream to each user, i.e., $|\mathbb{K}| = D$.

**B. Resource Allocation**

With the codebook indices and the scaled SINR values of all users, the base station schedules the users and computes the ZF precoder as described in the previous section. To do so, it calculates the SINR approximations based on the scaled versions thereof. It holds [10], [14]

$$\Gamma_K = \frac{M}{P_{\text{tx}}} A_K \text{diag} (\gamma_k^*)_{k \in \mathbb{K}}. \quad (19)$$

Then, it schedules the users according to Algorithm 1 where $R_k := \sum_{q \in \mathbb{K}} \log_2(1 + e_q^T \Gamma_k e_q)$. Please see [10], [14] for a detailed description of Algorithm 1.

**Algorithm 1 Resource allocation**

1. $\mathbb{K} \leftarrow \{\}$, $\hat{G}_K \leftarrow 0$
2. While $|\mathbb{K}| \leq M$
   1. $\mathbb{K}' \leftarrow \{k \in \{1, \ldots, K\} \setminus \mathbb{K} : \gamma_k^* \neq e_q^T \hat{G}_K \forall q\}$
   2. For $k \in \mathbb{K}'$
      1. $\hat{G}_K \leftarrow [\hat{G}_K ; g_k^T]$.
      2. Compute $P_{K \cup \{k\}}$ and $\Lambda_{K \cup \{k\}} \triangleright$ Eqs. (17) and (18), compute $\Gamma_{K \cup \{k\}} \triangleright$ Eqs. (19).
   3. End for
   4. $k' \leftarrow \arg \max_{k \in \mathbb{K} \cup \{k\}} R_{K \cup \{k\}}$
   5. If $R_{K \cup \{k'\}} > R_{K}$ then $\mathbb{K} \leftarrow \mathbb{K} \cup \{k'\}$
   6. Else exit
   7. End if
End while

Finally, the set $\mathbb{K}$ of scheduled users is used to compute the ZF precoder according to Eqs. (17) and (18).

**C. Simulation Results**

We investigate the proposed schemes in a MIMO Orthogonal Frequency Division Multiplex (OFDM) system with the parameters as given in Table I and assuming the typical urban macro-cell channel model of the WINNER project [16].

Figure 3 illustrates the performance difference between the CVQ schemes with Pseudo-Maximization (PM) and full maximization of the CQI indicator (approximate SINR), and compares them with the minimum Euclidean Distance (ED, minimum quantization error) approach. Here, we assume a random codebook with $B = 4$, where the elements of $\mathcal{C}$ are chosen from an isotropic distribution on the $M$-dimensional unit sphere, i.e., normalized versions of vectors with random entries according to $\mathcal{N}_C(0, 1)$. The maximal gain of the pseudo and full maximization schemes over the minimum Euclidean distance method seems
to be about 1.2 bits/s/Hz at 0 dB. Moreover, one can see that the pseudo-maximization scheme acceptably approaches the performance of the full maximization scheme. Indeed, the performance gap between the two is never more than about 0.7 bits/s/Hz. Surprisingly, pseudo-maximization and Euclidean distance minimization are both slightly superior to full maximization in the mid and high SNR regions. This is possible due to the fact that the SINR measure used as the cost criterion for maximization does not represent the exact SINR, but is rather an approximation of this quantity. It appears that Euclidean distance minimization is the best option at mid and high SNR, and that in that range the SINR pseudo-maximization achieves the same performance. As pseudo-maximization makes a choice between the codebook entry that maximizes channel magnitude and the one that minimizes the quantization angle (Euclidean distance), one can suspect that at high SNR the minimum quantization angle solution is chosen most of the time, such that no performance difference is noticeable with the scheme that only chooses the entry with the minimum Euclidean distance.

Remember that the pseudo-maximization scheme requires much less computational complexity than full maximization (see the discussion to that effect in Section III-B). Due to the above results, it also appears that pseudo-maximization always performs better or equivalently to quantization angle minimization, and that not much is gained by using full maximization (it can even cause slight performance degradation in certain SNR ranges). Therefore, the pseudo-maximization scheme seems like the preferred alternative to quantization angle minimization and produce significant sum rate gains with respect to this scheme in the low SNR region.

REFERENCES


