Low Complexity MMSE Vector Precoding Using Lattice Reduction for MIMO Systems

Feng Liu, Lingge Jiang, and Chen He
Department of Electronic Engineering
Shanghai Jiao Tong University
Shanghai 200240, China
{fengliu, lgjiang, chenhe}@sjtu.edu.cn

Abstract—In this paper, a lattice-reduction-aided (LRA) minimum mean square error (MMSE) vector precoding (VP) is proposed for multiple input multiple output (MIMO) systems. Three schemes are provided for the perturbation vector design by Babai’s approximation procedures based on lattice reduction method to reduce the complexity. Performance and complexity analysis are provided. Simulation results show that the proposed schemes significantly outperform the conventional MMSE Tomlinson-Harashima precoding (THP) and the zero forcing (ZF) VP. Compared with the MMSE VP via closest-point search, our LRA approach provides a simple alternative with little performance loss.

Keywords—MMSE vector precoding; lattice reduction; MIMO systems

I. INTRODUCTION

It has been shown that the multiple input multiple output (MIMO) systems can greatly increase the channel capacity and improve the system performance than the single antenna systems [1-3]. If the channel state information (CSI) is available at the transmitter, precoding technique can be implemented for further enhancement of system performance [4], [5]. Precoding is a transmit processing which is applicable to point to point channel and especially the broadcast channel where receiver cooperation is impossible [6].

The authors design a “vector perturbation technique” in [7], [8]. The key idea is to choose a vector of perturbation symbols to reshape the transmit symbols, which was presented in shaping without scrambling [9]. This technique is also called “vector precoding”. In fact it is a generalized Tomlinson-Harashima precoding (THP). In such schemes, the CSI is assumed at the transmitter to reshape the signals with a perturbation vector and only a simple modulo operation and a decision device are needed at the receiver. It is reported that VP can perform with full diversity. However, the perturbation vector is found via a closest-point search in lattice which has very high complexity.

In [10], the authors use lattice reduction method and Babai’s approximate closest-point solution [11] to reduce the complexity of VP along the line of [12] and [13]. Precoding/decoding schemes based on lattice reduction method are often referred to as lattice-reduction-aided (LRA) schemes.

With such an approximation, the complexity of VP is greatly reduced.

In fact, the above VP schemes are zero-forcing (ZF) approaches. It is well known that noise amplification of ZF is significant and harms the performance. On the other hand, the minimum mean square error (MMSE) approach offers the optimal tradeoff between noise amplification and residual interference in a mean square error (MSE) sense. MMSE approach based lattice reduction is used for decoding at the receiver in [14] and for precoding at the transmitter in [15]. MMSE VP is described in [16] and outperforms the ZF VP. However, the closest-point search in lattice makes it computationally intensive. In [17], the authors develop a new precoding scheme named scaled vector precoding (SVP) which divides the symbols into groups and then computes the perturbation vector in the group. However, the meaning of grouping is not well exploited and the optimizing of group number is not considered. Moreover, the diversity order of SVP lies between the diversity orders of THP and VP, i.e. full diversity is not always guaranteed.

In this paper, we propose a low complexity LRA MMSE VP for MIMO systems. With lattice reduction method and Babai’s approximation procedures, we design three low complexity LRA MMSE VP schemes. Simulation results show that the proposed algorithms outperform the conventional THP and the ZF VP (even the one doing closest-point search).

Notation: Throughout the paper, we denote vectors and matrices by lower and upper case bold letters, respectively. We use $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^i$, $\mathbb{R}\{\}$, and $\mathbb{I}\{\}$ for the transpose, pseudo inverse, inverse, the real part, and the imaginary part, respectively. The identity matrix is denoted as $\mathbf{I}$.

The rest of the paper is organized as follows. In section II, the system model is presented. In section III, the LRA MMSE VP schemes are proposed. Simulation results are provided in section IV. Section V concludes this paper.

II. SYSTEM MODEL

We consider a MIMO system with $M_T$ transmit and $M_R$ receive antennas. Block transmission with a length $N_B$ is used. There is a total power constraint $P_t$ for the
transmitted symbol block. Our derivations are restricted to frequency flat channels and perfect CSI is assumed at the transmitter. Note our approach can be directly extended to frequency selective channel and broadcast channel.

The signal model is given by

\[ y_c[k] = H_c x_c[k] + n_c[k] \quad k = 1, \ldots, N_u \]  

where \( y_c[k] \) is the \( k \) th \( M_u \times 1 \) complex symbol vector in the received block, \( H_c \) is the \( M_u \times M_r \) Rayleigh channel whose elements are independent identical distribution (i.i.d.) complex Gaussian with zero mean and unit variance, \( x_c[k] \) is the \( k \) th \( M_r \times 1 \) complex symbol vector in the transmitted block, \( n_c[k] \) is the corresponding \( M_r \times 1 \) complex vector of i.i.d. white Gaussian noise with variance \( \sigma_n^2 \). Without any confusion, we omit the index \([k]\) in the following description.

Employing a real-value notation, the signal model (1) can be expressed as

\[ y = H x + n \]  

where the real vectors are obtained from stacking the real and imaginary parts of the corresponding complex vectors as

\[
\begin{align*}
\Re\{y_c[k]\} & = \Re\{x_c[k]\} \\
\Im\{y_c[k]\} & = \Im\{x_c[k]\} \\
\Re\{n_c[k]\} & = \Re\{n[k]\} \\
\Im\{n_c[k]\} & = \Im\{n[k]\}
\end{align*}
\]

and the real channel matrix \( H \) is obtained as

\[ H = \begin{bmatrix}
\Re\{H_c\} & -\Im\{H_c\} \\
\Im\{H_c\} & \Re\{H_c\}
\end{bmatrix} \]

VP schemes generate the transmit symbols \( x \) by

\[ x = \beta F(s + p) \]  

where \( F \) is the precoding matrix, \( s \) is the real signal vector with a dimension \( K = 2M_u \), and \( p \) is the real perturbation vector which is a point in the integer lattice \( \alpha \mathbb{Z}^K \) with integer scalar \( \alpha \), and \( \beta \) is a power constraint factor.

In this paper, we use the square \( M \)-point Gray coded quadrature amplitude modulation (QAM) with the constellation \( \{ \pm 1/2, \ldots, \pm (\sqrt{M} - 1)/2 \}^2 \) for the complex model. The signal variance can be computed as \( \sigma_s^2 = (M - 1)/6 \). For the real value model the symbols of \( s \) are independently chosen from the corresponding \( A \)-ary amplitude-shift keying (ASK) \( (A = \sqrt{M}) \) with the constellation \( \{ \pm 1/2, \ldots, \pm (A-1)/2 \} \). The scalar is chosen as \( \alpha = A \).

III. PROPOSED SCHEMES

For ZF VP, the precoding matrix \( F \) is chosen as the right pseudo inverse of the channel matrix

\[ F_{ZF} = H_{ZF}^H = H^H (HH^H)^{-1} \]  

and the perturbation vector \( p \) is found by minimizing the transmit power as the following searching rule

\[ p_{ZF} = \arg \min_{p \in \mathbb{Z}^K} \| H_{ZF} (s + p) \|^2. \]  

It is well known that MMSE approach makes an optimal tradeoff between the residual interference and noise to reduce the noise amplification effect. For MMSE VP [16], the precoding matrix is a MMSE version inverse of channel matrix

\[ F_{MMSE} = H_{MMSE}^H \triangleq H^H (HH^H + \sigma_n^2 / \sigma_s^2 I)^{-1} \]

and the perturbation vector is found by the searching rule

\[ p_{MMSE} = \arg \min_{p \in \mathbb{Z}^K} \| L(s + p) \|^2 \]  

where \( L \) is a \( K \times K \) matrix obtained from the Cholesky factorization \( (HH^H + \sigma_n^2 / \sigma_s^2 I)^{-1} = L^T L \)

At the receiver, modulo operation \( \mod \alpha \) is performed to eliminate the perturbation vector. Finally, the detector simply quantizes the symbols to the constellation for estimation.

The optimal solution of finding the optimal vector \( p \) is the maximum-likelihood (ML) method by searching over the whole integer lattice. However, because the lattice is infinite, ML method is impossible to implement. A feasible alternative is the closest-point search algorithm (sphere encoding), which restricts the search space to a sphere. But the computational complexity is still very high in comparison with the simple and suboptimal linear and successive interference cancellation (SIC) processing, which are just done by Babai’s approximation. Here we use the lattice reduction method and Babai’s closest-point approximation to reduce the complexity of MMSE VP.

Firstly, we apply the polynomial-time LLL (Lenstra, Lenstra, Lovasz) algorithm [18] for lattice reduction to \( L \) as

\[ L_{\text{red}} = LT \]  

where \( L_{\text{red}} \) is the LLL-reduced basis with approximately orthogonal columns, and \( T \) is a matrix with integer entries and determinant \( \pm 1 \).

Then, with Babai’s two polynomial-time algorithms, called rounding off procedure and nearest plane procedure, respectively, we propose three schemes to approximate the perturbation vector design by linear and SIC processing. They have polynomial complexity and avoid the high complexity sphere encoding algorithm. To use Babai’s algorithms, the signal needs a shift to integer grid. We omit its description in the following discussion.

A. Scheme 1

With the rounding off procedure, linear processing is performed. The perturbation vector in (7) is approximated as

\[ p_1 = -T Q_{\alpha \mathbb{Z}^K} \{ T^T s \} \]  

where \( Q_{\alpha \mathbb{Z}^K} \{ \} \) denotes component-wise rounding of the vector to the scaled integer lattice \( \alpha \mathbb{Z}^K \).
Scheme 2

With the nearest plane procedure, SIC processing is performed. By the QR factorization of $L_{\text{red}}$, we have

$$QR = L_{\text{red}}$$  \hspace{1cm} (10)

where $Q$ is a unitary matrix, and $R$ is an upper-triangular matrix with elements $[r_{ij}]$.

Let $D = \text{diag}(1/r_1, \ldots, 1/r_K)$ be diagonal with the inverse of the diagonal elements of $R$. We construct an upper-triangular matrix with unit diagonal elements

$$B = DR.$$  \hspace{1cm} (11)

Now with the nearest plane procedure we set

$$q = -DQ^\dagger Ls$$  \hspace{1cm} (12)

and compute $\tilde{q}_k = q_k$, and for $i = K - 1, \ldots, 1$

$$\tilde{q}_i = Q_{az}\left\{q_i - \sum_{j=i+1}^{K} b_{ij}\tilde{q}_j\right\}.$$  \hspace{1cm} (13)

Let $\tilde{q} = (\tilde{q}_1, \ldots, \tilde{q}_K)^\dagger$. The approximated perturbation vector is

$$p_2 = T\tilde{q}. \hspace{1cm} (14)$$

C. Scheme 3

Considering the SIC with optimal processing order, we can apply the V-BLAST algorithm [19] to $L_{\text{red}}$

$$WL_{\text{red}}P = B$$  \hspace{1cm} (15)

where $W$ is a matrix with orthogonal rows, $P$ is a permutation matrix indicating the optimal order, and $B = [b_{ij}]$ is a lower triangular matrix with unit diagonal. Now with the nearest plane procedure, we set

$$q = -WLs$$  \hspace{1cm} (16)

and compute $\tilde{q}_i = q_i$, and for $i = 2, \ldots, K$

$$\tilde{q}_i = Q_{az}\left\{q_i - \sum_{j=1}^{i-1} b_{ij}\tilde{q}_j\right\}.$$  \hspace{1cm} (17)

Thus, we obtain the approximated perturbation vector

$$p_3 = TP\tilde{q}.$$  \hspace{1cm} (18)

The above proposed LRA MMSE VP schemes are shown in Fig. 1. Note the factor $\beta$ is compensated at the receiver by automatic gain control. Our main contribution is to design the approximation algorithms to reduce the complexity of finding the perturbation vector for MMSE VP.

---

Figure 1. Illustration of the three proposed schemes. Top to bottom: Scheme 1 (LRA MMSE VP by linear approximation); Scheme 2 (LRA MMSE VP by SIC approximation via QR); Scheme 3 (LRA MMSE VP by ordered SIC approximation via V-BLAST).
D. Performance analysis

The difference between our design and the LRA ZF VP lies in both the precoding matrix and the perturbation vector design. In our design, the perturbation vector obtained by Babai’s approximation and precoding matrix will approximately minimize the MSE by offering a tradeoff between noise amplification and residual interference.

Let the singular value decomposition (SVD)

\[ H = U \Sigma V^T \]  

where \( U \) and \( V \) are unitary matrices and \( \Sigma \) is diagonal with non-negative singular values \([\sigma_i]\) in descending order.

Define

\[ E_{ZF} \triangleq \|H_{ZF}^T(s+p)\|^2 \]

\[ = (s+p)^T(HH^\dagger)^{-1}(s+p) \]

\[ = (s+p)^T U \Sigma^{-2} U^T (s+p) \]  

\[ = \sum_{j=1}^{\kappa} \frac{1}{\sigma_j} |u_j^T(s+p)|^2 \]  

and

\[ E_{MMSE} \triangleq \|L(s+p)\|^2 \]

\[ = (s+p)^T (HH^\dagger + \sigma_n^2 / \sigma_s^2 I)^{-1}(s+p) \]

\[ = (s+p)^T U (\Sigma^2 + \sigma_n^2 / \sigma_s^2 I)^{-1} U^T (s+p) \]  

\[ = \sum_{j=1}^{\kappa} \frac{1}{\sigma_j^2 + \sigma_n^2 / \sigma_s^2} |u_j^T(s+p)|^2 \]  

as the optimizing goals of ZF and MMSE searching rules, respectively, which represent the MSE [16]. It is clear that

\[ E_{MMSE} < E_{ZF} \]  

holds for the same signal and perturbation vector. Thus the MSE of MMSE VP will be smaller than that of ZF VP, which indicates MMSE VP will outperform ZF VP. This result holds for VP with lattice reduction method. On the other hand, we can see the noise variance subtly affects the choosing of perturbation vector. The ZF approach omits the noise item, which results in a sub-optimal lattice point, i.e. a sub-optimal perturbation vector.

Scheme 1 does the linear approximation of the perturbation vector. With SIC approximation in the other two schemes, better performance can be obtained than Scheme 1. However, due to the approximation processing, there is some performance loss for the LRA MMSE VP schemes compared with the MMSE VP via closest point search.

E. Complexity analysis

We use the big O notation to describe the computation complexity of the related algorithms. If only the multiplication operation for the approximation of the perturbation vector is counted, the complexity of proposed schemes are computed and shown in Table 1.

### TABLE I. COMPLEXITY COMPARISON OF PROPOSED SCHEMES

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Computation Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme 1</td>
<td>(O(\max{K^1, K^2 N_n}))</td>
</tr>
<tr>
<td>Scheme 2</td>
<td>(O(\max{K^3, K^2 N_n}))</td>
</tr>
<tr>
<td>Scheme 3</td>
<td>(O(\max{K^4, K^2 N_n}))</td>
</tr>
</tbody>
</table>

We see the proposed schemes have polynomial complexity. Schematic 1 and Scheme 2 show the same complexity, since the matrix inverse in Scheme 1 and the QR decomposition in Scheme 2 have equal complexity of \(O(K^3)\). However, the V-BLAST algorithm in Scheme 3 has a higher complexity of \(O(K^4)\), which needs more computation to obtain the optimal processing order for SIC. Both linear and SIC processing have complexity of \(O(K^2 N_n)\) for the transmit symbol block at each channel instance.

Compared with the conventional THP, the complexity improvement of LRA precoding comes from the LLL algorithm which has complexity \(O(K^4)\). From Table 1 we know that if the block length \(N_n\) is large enough to dominate the complexity, the complexity improvement of LLL algorithm will be negligible. In this case, the proposed LRA MMSE VP schemes have the same complexity as the conventional THP. Therefore, our design with lattice reduction approximation greatly reduces the complexity compared with the MMSE VP via closest-point search.

IV. SIMULATION RESULTS

The simulation is set for a typical \(M_t = M_r = 4\) MIMO system. The performance is measured in terms of the uncoded bit error rate (BER) versus the average received signal-to-noise ratio (SNR) which is defined as \(E_i / N_0 = M_t \sigma_s^2 / R_m \sigma_n^2\), where \(R_m\) is the modulation rate, i.e. bits per modulated symbol. The three proposed schemes are labeled as “LRA MMSE VP by linear”, “LRA MMSE VP by QR” and “LRA MMSE VP by VB”, respectively. For comparison, the performance results of conventional MMSE THP (by QR and by VB), the LRA ZF VP schemes in [10] (labeled as “LRA ZF VP by linear”, “LRA ZF VP by VB”) and the VP schemes by closest-point search method (labeled as “ZF VP by search” and “MMSE VP by search”) are provided. The numerical results using 4-QAM, 16-QAM and 64-QAM are depicted in Fig. 2, Fig. 3, and Fig. 4, respectively.

The BER curves show that our LRA MMSE VP schemes significantly outperform the corresponding LRA ZF VP and conventional MMSE THP schemes. With rounding off procedure, the LRA MMSE VP achieves an improvement of about 2.5dB, 2dB and 1dB in comparison with the LRA ZF VP at BER of \(10^{-5}\) for 4-QAM, 16-QAM and 64-QAM, respectively. With the nearest plane procedure, the gain is about 3dB, 2.5dB and 1dB for 4-QAM, 16-QAM and 64-QAM, respectively. The LRA MMSE VP with the nearest plane
procedure even significantly outperforms the ZF VP which does the closest-point search. The “LRA MMSE VP by QR” scheme performs nearly the same as the “LRA MMSE VP by VB”, while the conventional MMSE THP by VB significantly outperforms the conventional MMSE THP by QR. Therefore, the SIC processing ordering for LRA precoding does not greatly affect the performance, while the conventional MMSE THP significantly benefits from optimal processing ordering. We contribute it to the application of the LLL algorithm for lattice reduction by greatly eliminates the interference among the MIMO channels.

Furthermore, all ZF VP schemes show a diversity order of 4, while the conventional MMSE THP only achieves a diversity order of 1. Interestingly, all MMSE VP schemes show a diversity order of 5 which is larger than that of the ZF VP. Here we contribute it to the MMSE approach, which optimally balances the noise amplification and residual interference, and the use of real-value model, which expands the channel dimension and leads to some additional degrees of freedom.

Compared with MMSE VP via closest-point search, the first scheme has a power efficiency loss of about 3.5dB, 2dB and 1.8dB at BER of $10^{-5}$ for 4-QAM, 16-QAM and 64-QAM, respectively. For the other two schemes, the loss is about 1.5dB for 4-QAM at BER of $10^{-5}$. When larger constellation of modulation is used, the power efficiency loss reduces. In our simulation, the loss is less than 0.5dB at BER of $10^{-5}$ for 16-QAM and 64-QAM. If block transmission with enough large size is assumed for each channel instance, the complexity of the LRA MMSE VP would reduce to that of conventional THP. Therefore, the proposed schemes (especially Scheme 2 and Scheme 3) provide a simple and effective alternative for the MMSE VP doing lattice search at the cost of tolerable performance loss.

**V. CONCLUSIONS**

We presented a low complexity MMSE VP with lattice reduction approximation. The LRA MMSE VP greatly reduces the noise amplification compared with the LRA ZF VP. Based on the polynomial complexity LLL algorithm, we used Babai’s procedures to perform the linear and SIC approximation for the perturbation vector design. QR factorization and V-BLAST algorithm were utilized to develop two SIC approximation schemes without/with optimal processing order. Simulation results verify our analysis and design. Some conclusions can be drawn as follows: (1) the proposed schemes (especially the SIC approximation) significantly outperform the (LRA) ZF VP and conventional MMSE THP with polynomial complexity; (2) if the block length is large enough to dominate the complexity of the scheme, the proposed LRA MMSE VP schemes share the same complexity as the conventional THP; (3) compared with the MMSE VP which does high complexity closest-point search for the perturbation vector, the performance loss of the proposed schemes by lattice reduction approximation is very small (less than 0.5dB for the SIC approximation); (4) higher diversity order is obtained by the use of MMSE approach and real model than ZF VP; (5) the processing ordering is nearly useless for LLL-based LRA VP schemes, while conventional THP greatly benefits from it.
REFERENCES