Communication Over MIMO Broadcast Channels Using Lattice-Basis Reduction

Mahmoud Taherzadeh, Amin Mobasher and Amir. K. Khandani
Coding & Signal Transmission Laboratory(www.cst.uwaterloo.ca)
Dept. of Elec. and Comp. Eng., University of Waterloo, Waterloo, ON, Canada, N2L 3G1
e-mail: {taherzad, amin, khandani}@cst.uwaterloo.ca, Tel: 519-8848552, Fax: 519-8884338

September 27, 2004

Abstract

A simple scheme for communications in MIMO broadcast channels is introduced which adopted the lattice reduction techniques to improve the naive channel inversion method. Lattice basis reduction helps us to reduce the average transmitted power by modifying the region which consists of the constellation points. Simulation results show that the proposed scheme performs well and as compared to the complex methods (such as the perturbation method [1]) has a negligible loss. Also, using our scheme the method in [1] is modified which outperforms the perturbation method while it has a lower complexity.

1 Introduction

In the recent years, MIMO communications over multiple-antenna channels has attracted many researchers. In the last decade, the main interest has been on the point-to-point MIMO communications. However, recently, new information theoretic results [2], [3], [4], [5], have shown that also in multiuser MIMO systems we can exploit many of advantages of multiple-antenna systems. It has been shown that in a MIMO broadcast system with K users (with single antenna) and a transmitter with M antennas, the sum-capacity grows linearly with the minimum of M and K.

To achieve the sum capacity, some information theoretic schemes, based on dirty-paper coding, are introduced. Some methods, such as using nested lattices [6] are introduced as practical techniques to achieve the sum-capacity promised by dirty-paper coding. However, these methods are not easy to implement.

In [1], the authors have introduced a *vector perturbation technique* which has a good performance in terms of symbol error rate. Nonetheless, this technique requires a lattice-decoder which is an NP-hard problem. In [7], The authors have used lattice-basis reduction to approximate the closest lattice point (using Babai approximation).

In this paper, we present a transmission technique, based on the lattice-basis reduction, for the MIMO broadcast channels. Instead of approximating the closest lattice point in the perturbation problem, we use the lattice-basis reduction to reduce the average transmitted power by reducing the second moment of the fundamental region of the transmitted lattice. We show that the proposed method has almost the same performance as [1] with a much smaller complexity. As compared to [7], the proposed code offer a better performance. Also, the proposed viewpoint helps us to expand our approach for unequal-rate transmissions.

2 System Model

We consider a multiple-antenna system with an access point with M transmit antennas and K users. In this paper, we focus on the case that each user has one receive antenna. If we consider $\mathbf{y} = [y_1, ..., y_K]^T$, $\mathbf{x} = [x_1, ..., x_K]^T$, $\mathbf{w} = [w_1, ..., w_K]^T$ and the $K \times M$ matrix $\mathbf{H}$, respectively, as the received signal, the transmitted signal, the noise vector and the channel matrix, we have the following matrix equation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}. \tag{1}$$

The channel is assumed to be Raleigh and the noise is Gaussian, i.e. the elements of $\mathbf{H}$ are i.i.d with the zero-mean unit-variance complex Gaussian distribution. Also, we have the power constraint on the transmitted signal, $\mathrm{E}\|x\|^2 = 1$. The power of the additive noise is σ^2 per antenna, i.e. $\mathrm{E}\|w\|^2 = K\sigma^2$. Therefore, the signal to noise ratio (SNR) is defined as $\rho = \frac{1}{\sigma^2}$.

In the decoding of the received signal, the users can not cooperate with each other. To resolve this problem, we can construct the transmitted signals such that the interference is cancelled in the receiver, i.e. different users see independent signals (only their data which is added by the additive noise). The simplest method is the channel inversion:

$$\mathbf{s} = \mathbf{H}^{+}\mathbf{u}, \tag{2}$$

where $\mathbf{H}^{+} = \mathbf{H}^{*}(\mathbf{H}\mathbf{H}^{*})^{-1}$, $\mathbf{s}$ is the transmitted signal before the normalization, and $\mathbf{u}$ is the data vector, i.e. u_i is the data for the i'th user which can be a QAM point. When $M = K$, the transmitted signal becomes

$$\mathbf{s} = \mathbf{H}^{-1}\mathbf{u}. \tag{3}$$

The problem arises when $\mathbf{H}$ is poorly conditioned and $\|\mathbf{s}\|$ becomes very large. This results high power consumption or very small received signals (if we normalized the transmitted signal). To reduce the average power of $\mathbf{s}$, we can use the lattice reduction methods.

3 Lattice-Basis Reduction

Lattice structures have been used frequently in different communication applications such as quantization or Multiple-Input Multiple-Output (MIMO) decoding. A real lattice Λ is a discrete set of M-D vectors in real Euclidean M-space $\mathcal{R}^M$ that forms a group under ordinary vector addition. Every lattice Λ is generated by the integer linear combinations of some set of linearly independent vectors $\mathbf{b}_1, \cdots, \mathbf{b}_N \in \Lambda$, where the integer $N (\leq M)$ is called the dimension of the lattice Λ (Here we assume that N=M). The set of vectors $\mathbf{b}_1, \cdots, \mathbf{b}_N$ is called a basis of Λ, and the matrix $B = (\mathbf{b}_1, \cdots, \mathbf{b}_N)$ which has the basis vectors as its columns is called the basis matrix (or generator matrix) of Λ.

The basis for representing a lattice is not unique. Any ordered basis of Λ, say $\mathbf{b}_1, \cdots, \mathbf{b}_N$, can be assigned a set of Gram-Schmidt vectors, $\hat{\mathbf{b}}_1, \cdots, \hat{\mathbf{b}}_N$, which are mutually orthogonal and can be computed using the following recursion:

$$\begin{aligned} \hat{\mathbf{b}}_i &= \mathbf{b}_i - \textstyle\sum_{j=1}^{i-1} \mu_{ij} \hat{\mathbf{b}}_j, \text{ with} \\ \mu_{ij} &= \frac{< \mathbf{b}_i, \hat{\mathbf{b}}_j >}{< \mathbf{b}_j, \hat{\mathbf{b}}_j >} \end{aligned} \tag{4}$$

Usually a basis consisting relatively short and nearly orthogonal vectors is desirable. The procedure of finding such a basis for a lattice is called *Lattice Basis Reduction.* Several distinct notions of reduction have been studied, including those associated to the names Minkowski, Korkin-Zolotarev, and more recently Lenstra-Lenstra and Lovasz (LLL) reduced basis, which can be computed in polynomial time.
A basis $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ is *Minkowski-Reduced Basis* [8] if

- $\mathbf{b}_1$ is the shortest nonzero vector in the lattice Λ, and
- For each k=2,...,N, $\mathbf{b}_k$ is the shortest nonzero vector in Λ such that $\{\mathbf{b}_1, \cdots, \mathbf{b}_k\}$ may be extended to a basis of Λ.

Finding Minkowski reduced basis is equivalent to finding the shortest vector in the lattice and this problem by itself is NP-hard. Thus, there is no polynomial time algorithm for this reduction method. *Korkin-Zolotarev (KZ)* reduction is a variant of Minkowski reduction that has proven very useful for algorithmic as well as mathematic methods.
A basis $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ is *KZ Reduced basis* [9] if

- $\mathbf{b}_1$ is the shortest nonzero vector in the lattice Λ, and
- For each k=2,...,N, $\mathbf{b}_k$ is the shortest nonzero vector in Λ_i which is obtained by projecting Λ to the subsapce of $\mathcal{R}^N$ perpendicular to $\{\mathbf{b}_1, \cdots, \mathbf{b}_{k-1}\}$.

In the definition of Minkowski reduction, successive basis vectors are added to the lattice basis only if b_i is the shortest vector in the lattice which will allow the basis to be extended. In KZ reduction, though, successive basis vectors are chosen based on their length in the orthogonal complement of the space spanned by the previous basis vectors $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$
It can be shown that for each lattice there is at least one KZ basis [10]. However, there is no polynomial time algorithm known for KZ reduction.
A basis $\{\mathbf{b}_1, \cdots, \mathbf{b}_N\}$ is *LLL Reduced Basis* [11]

- $|mu_{ij} \leq \frac{1}{2}$ for $1 \leq i < j \leq N$, and
- $|\hat{b}_i|^2 \leq \dfrac{4}{3} |\hat{b}_{i+1} + \mu_{i+1,i} \hat{b}_i|^2$

It is shown that LLL reduced basis algorithm produces lattice bases with relatively short vectors in a polynomial time computational algorithm [11]. However, any polynomial time lattice basis reduction algorithm we use will not be able to satisfy the strict conditions of Minkowski or KZ reduction. LLL reduced basis has extended applications in several contexts due to its polynomial time algorithm.

4 Proposed Approach

Assume that the data for different users, u_i, is selected from the points of QAM constellations. Now, the data vector $\mathbf{u}$ is a point in the Cartesian product of these QAM constellations. Therefore, the data region can be considered as a $2K$ dimensional hypercube[1]. At the transmitter, when we use the channel inversion technique, the transmitted signal is a point inside a parallelotope whose sides are in the direction of the columns of $\mathbf{H}^{-1}$. If the data is a point from the integer lattice $\mathcal{Z}^{2K}$, the transmitted signal is a point in the lattice generated by $\mathbf{H}^{-1}$. However, due to the lack of cooperation among the users, we need to sustain the data in the integer lattice $\mathcal{Z}^{2K}$ (to avoid the interference). Therefore, we need to change the region of the transmitted signals without changing the underlying lattice. Thus, lattice basis reduction is the solution.

When we use the continuous approximation (which is appropriate for large constellations), the average power of the transmitted signal is approximated by the second moment of the transmitted region. When we assume equal rates for the users e.g. rate r bits per user, the data is inside a hypercube whose sides are equal to $a = 2^{r/2}$. Therefore, the transmitted region is the scaled version of the fundamental region of the transmitted lattice (corresponding to its basis) with the scaling factor a. The second moment of this region is proportional to the sum of squared norms of the basis vectors (see Appendix). Therefore, we should try to find a basis reduction method which minimizes the sum of squared norms of the basis vectors. Among the famous reduction algorithms, the Minkowsky reduction can be considered as a good greedy algorithm for our problem. Indeed, the Minkowsky algorithm is the successively optimum solution because in each step, it finds the smallest vector. However, the complexity of the Minkowsky reduction is equal to the complexity of the lattice-decoding problem. Therefore, we use LLL reduction algorithm which is a suboptimum solution with a polynomial complexity.

Assume that $\mathbf{B} = \mathbf{H}^{-1}\mathbf{U}$ is the reduced basis for the lattice obtained by $\mathbf{H}^{-1}$ and $\mathbf{u}$ is the data vector. Now, we use $\mathbf{x} = \mathbf{B}\mathbf{u}' = \mathbf{H}^{-1}\mathbf{U}\mathbf{u}'$ as the transmitted signal where

$$\mathbf{u}' = \mathbf{U}^{-1}\mathbf{u} \mod a. \tag{5}$$

At the receiver, we use modulo operation to find the actual data:

$$\begin{aligned} \mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{H}^{-1}\mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod a) + \mathbf{n} \\ = \mathbf{U}(\mathbf{U}^{-1}\mathbf{u} \mod a) + \mathbf{n} = \mathbf{U}\mathbf{U}^{-1}\mathbf{u} \mod a + \mathbf{n} = \mathbf{u} \mod a + \mathbf{n}. \end{aligned} \tag{6}$$

In practice, most of the times, we are interested in using points of the lattice with odd coordinates. When the data vector $\mathbf{u}$ consists of odd integers, if we use the lattice-basis reduction, instead of points with all-odd coordinates in the new basis, we may have points which have some even coordinates i.e. $\mathbf{U}^{-1}\mathbf{u}$ has some even elements. In this case, the constellation is not symmetric in respect to the origin. Therefore, we can reduce the transmitted power by shifting the center of the constellation to the origin. The translation vector is equal to $(\mathbf{U}^{-1}[1+i, 1+i, \cdots, 1+i]^T + [1+i, 1+i, \cdots, 1+i]^T) \mod 2$. When we use this shifted version of the constellation, we must send the translation vector to the users (2 bits for each of them), at the beginning of the block. However, compared to the whole block of data, these two bits are negligible. Besides, in all scenarios of the MIMO broadcast channels with CSI, we should send a training sequence for the users, at the beginning of the transmission, to estimate the channel coefficients.

[1] Each QAM constellation is a two dimensional square.

In the proposed method at the beginning of each fading block we reduce the lattice obtained by $\mathbf{H}^{-1}$ and during this block the transmitted signal is obtained by equation (5), which has a complexity in the order of a matrix multiplication and a modulo operation. Therefore, the complexity of the proposed method is like that in the channel inversion method; however, as we will show the performance of this method is near the performance of the perturbation method in [1].

We can even further improve the performance of the proposed method. After reducing the inversion of the channel matrix and obtaining the parity bits at the beginning of each fading block, the closest point to the signal computed in equation (5) can be found using the sphere decoder. Then, the transmitted signal is obtained using this point. For example for a lattice with odd coordinates, the transmitted signal can be obtained by

$$\mathbf{x} = \mathbf{B}\left(\mathbf{u}' + \mathbf{l} + \mathbf{u}_{par}\right), \tag{7}$$

where $\mathbf{u}_{par}$ is the parity vector, which is computed for users at the beginning of the fading block, and $\mathbf{l}$ is an even integer vector such that the vector $\mathbf{x}$ has the minimum power.

This method can be considered as *modified perturbation method*. It outperforms the perturbation method in [1]; however, it has a lower complexity because of employing the reduced matrix $\mathbf{B}$ instead of the matrix $\mathbf{H}^{-1}$ in the sphere decoder procedure.

5 Unequal-Rate Transmission

In the previous section, we consider the case that the transmission rates for different users are equal. If we are interested in sum-rate, instead of individual rates, we can improve the performance of the proposed method to reduce the probability of error by assigning unequal rates to the users.

In this section, we assume that the sum-rate is fixed (rather than individual rates) and we want to reduce the average transmitted power. To simplify the analysis, we use continuous approximation, which is a reasonable approximation for high rates.

Assuming continuous approximation, the sum-rate is proportional to the volume of the lattice with the basis $\mathbf{B}$ and the average power is proportional to $\sum_{i=1}^{M} |\mathbf{b}_i|^2 = \mathrm{tr}\mathbf{B}\mathbf{B}^*$ (see Appendix). The goal is minimizing the average power whilst the sum-rate is constant. Now, if there is not any constraint on individual rates, we can use another lattice generated by $\mathbf{B}'$ with the same volume, where its basis vectors are scaled versions of the columns of $\mathbf{B}$. Therefore, we can use $\mathbf{B}' = \mathbf{B}\mathbf{D}$ (where $\mathbf{D}$ is a unit determinant diagonal matrix) instead of $\mathbf{B}$. When $\mathbf{B}$ is fixed, the multiplication of the squared norm of the new basis vectors is constant:

$$\begin{aligned} |\mathbf{b}'_1|^2|\mathbf{b}'_2|^2...|\mathbf{b}'_M|^2 &= \left(|\mathbf{b}_1|^2|\mathbf{b}_2|^2...|\mathbf{b}_M|^2\right)\det\mathbf{D} \\ &= |\mathbf{b}_1|^2|\mathbf{b}_2|^2...|\mathbf{b}_M|^2 = \mathit{constant}; \end{aligned} \tag{8}$$

however, the average power corresponding to the new lattice basis must be minimized. According to the arithmetic-geometric mean inequality, $\sum_{i=1}^{M} |\mathbf{b}'_i|^2 = \mathrm{tr}\mathbf{B}'\mathbf{B}'^*$ is minimized iff

$$|\mathbf{b}'_1| = |\mathbf{b}'_2| = ... = |\mathbf{b}'_M|. \tag{9}$$

Having the matrix $\mathbf{B}$, the elements of the diagonal matrix $\mathbf{D}$ can be found using the equation (9). For the choice of the basis reduction, we should find $\mathbf{B}$ so that $|\mathbf{b}_1|^2|\mathbf{b}_2|^2...|\mathbf{b}_M|^2$ is minimized. Now, $\det\mathbf{B} = \det\mathbf{H}^{-1}$ is fixed. Thus, the best basis reduction is the reduction which maximizes $\frac{|\det\mathbf{B}|}{(|\mathbf{b}_1||\mathbf{b}_2|...|\mathbf{b}_M|)}$ or in the other word, minimizes the orthogonal deficiency[2]. Instead of a naive maximization for $\frac{|\det \mathbf{B}|}{(|\mathbf{b}_1||\mathbf{b}_2|...|\mathbf{b}_M|)}$ we can use a recursive algorithm. The first vector $\mathbf{b}_1$ is selected as the shortest nonzero vector in the lattice generated by $\mathbf{H}^{-1}$. Having the $M \times k$ matrix $\mathbf{B}_k = [\mathbf{b}_1, \cdots, \mathbf{b}_k]$, we can find the $(k+1)$th vector by

$$\max_{\mathbf{b}_{k+1}} \frac{\det \left(\mathbf{B}_{k+1}^* \mathbf{B}_{k+1}\right)}{|\mathbf{b}_1|^2 |\mathbf{b}_2|^2 ... |\mathbf{b}_{k+1}|^2}. \tag{10}$$

Theorem 1 *KZ reduction can be seen as a greedy algorithm for finding the proper reduced basis in the sense of maximization of equation (10) [12].*

6 Diversity and Outage Probability

When we have the channel-state information at the transmitter, if there is no assumption on the transmission rates, the outage probability is not meaningful. However, when we consider fixed rates $R_1, ..., R_K$ for different users, we can define the outage probability P_{out} as the probability that the point $(R_1, ..., R_K)$ is outside of the capacity region.

Theorem 2 *For a MIMO broadcast system with M transmit antennas and M single-antenna receivers and fixed rates $R_1, ..., R_K$ [12],*

$$\lim_{\rho \to \infty} \frac{-\log P_{out}}{\log \rho} \leq M.$$

We can also define the diversity gain of a MIMO broadcast constellation as the $\lim_{\rho \to \infty} \dfrac{\log P_e}{\log \rho}$ where P_e is the probability of error. Based on theorem 2, the maximum achievable diversity is M. We show that the proposed method (based on lattice-basis reduction) achieves the maximum diversity.

Lemma 1 *Consider $\boldsymbol{B} = [\boldsymbol{b}_1 ... \boldsymbol{b}_M]$ as an $M \times M$ matrix, with the orthogonality defect oth, and $(\boldsymbol{B}^{-1})^* = [\boldsymbol{a}_1 ... \boldsymbol{a}_M]$ as the Hermitian of its inverse. Then [12],*

$$\max\{|\boldsymbol{b}_1|, ..., |\boldsymbol{b}_M|\} \leq \frac{1}{(1 - oth). \min\{|\boldsymbol{a}_1|, ..., |\boldsymbol{a}_M|\}}. \tag{11}$$

Lemma 2 *Consider $\boldsymbol{B} = [\boldsymbol{b}_1 ... \boldsymbol{b}_M]$ as a reduced basis (Minkowski, KZ or LLL) for the lattice generated by $\boldsymbol{H}^{-1}$ and $d_{\boldsymbol{H}^*}$ as the minimum distance of the lattice genereted by $\boldsymbol{H}^*$. Then, there is a constant c (independent of $\boldsymbol{H}$) such that [12],*

$$\max\{|b_1|, ..., |b_M|\} \leq \frac{c}{d_{\boldsymbol{H}}}.$$

Lemma 3 *Assume that the entries of the $M \times M$ matrix $\boldsymbol{H}$ has independent complex Gaussian distributions with zero mean and unit variance and consider d_H as the minimum distance of the lattice generated by $\boldsymbol{H}$. Then, there is a constant C such that [12],*

$$Prob\{d_{\boldsymbol{H}} \leq \epsilon\} \leq C\epsilon^M.$$

Theorem 3 *For a MIMO broadcast system with M transmit antennas and M single-antenna receivers and fixed rates $R_1, ..., R_K$, When we use the lattice-basis-reduction method [12],*

$$\lim_{\rho \to \infty} \frac{-\log P_e}{\log \rho} = M.$$

[2] orthogonal deficiency is defined as $1 - \frac{|\det \mathbf{B}|}{(|\mathbf{b}_1||\mathbf{b}_2|...|\mathbf{b}_M|)}$

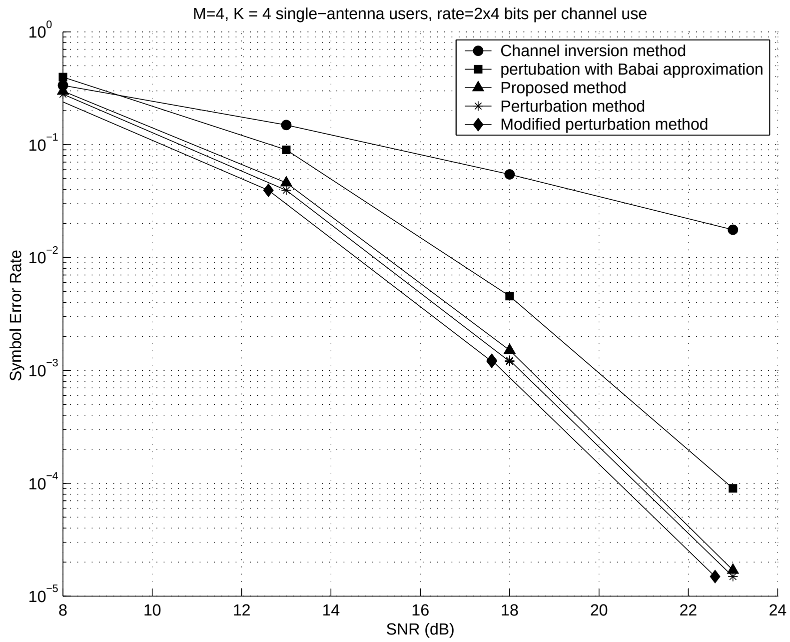

Figure 1: Symbol Error Rate of the proposed schemes, the perturbation scheme [1], and the naive channel inversion approach for $M = 4$ transmit antennas and $K = 4$ single-antenna receivers with the rate $R = 2$ bits per channel use per user.

7 Simulation Results

Figure 1 presents the simulation results for the performance of the proposed schemes, the perturbation scheme [1], and the naive channel inversion approach. The number of the transmit antennas is $M = 4$ and there are $K = 4$ single-antenna users in the system. The overall transmission rate is 16 bits per channel use, where 4 bits are assigned for each user.

We see that by using the proposed reduction-based schemes, we can achieve the full diversity, with a low complexity. Also, as compared to the perturbation scheme, we have a very little loss in the performance (about 0.2 dB). Also, compared to the approximated perturbation method [7], we have about 1.5 dB improvement by sending the parity bits at the beginning of the transmission. Also, the modified perturbation method (with sending two parity bits for each user) has around 0.3 dB improvement compared to the perturbation method while it has a lower complexity.

Figure 2. compares the performances of equal-rate and unequal rate transmission using lattice-basis reduction. In the both cases, the sum-rate is 8 bits per channel use. We see that by eliminating the equal-rate constraint, we can considerably improve the performance (especially, for high rates). In fact, the diversity gains for the equal-rate and the unequal-rate methods are respectively M and M^2.

8 Conclusion

A simple scheme for communications in MIMO broadcast channels is introduced which adopted the lattice reduction techniques to improve the naive channel inversion method. Lattice basis reduction helps us to reduce the average transmitted power by modifying the region which consists of the constellation points. Simulation results show that the

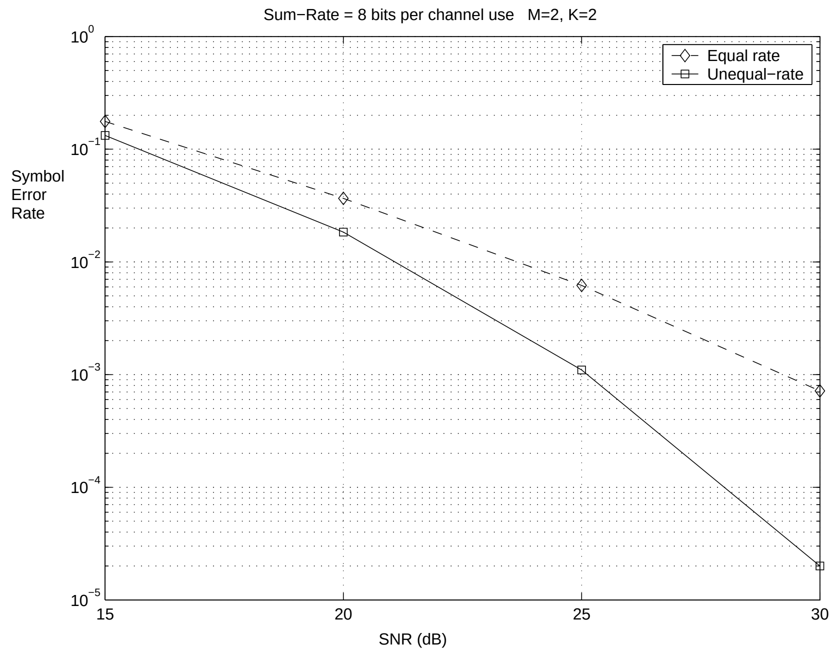

Figure 2: Performance comparison between the equal-rate and the unequal rate transmission for $M = 2$ transmit antennas and $K = 2$ single-antenna receivers with the sum-rate $R_{sum} = 8$ bits per channel use.

proposed scheme performs well and as compared to the complex methods (such as the perturbation method [1]) has a negligible loss. Also, the proposed modified perturbation method outperforms the perturbation method in [1] while it has a lower complexity.

Appendix A

In this Appendix, we compute the second moment of a parallelotope whose centroid is the origin and its sides are in the direction of the basis of the lattice.

Assume that $\mathcal{A}$ is the mentioned parallelotope with N dimensions and X is its second moment. Now the second moment of $\frac{1}{2}\mathcal{A}$ is $\left(\frac{1}{2}\right)^{N+2} X$. The parallelotope $\mathcal{A}$ can be considered as the union of 2^N smaller parallelotope which are constructed by $\pm\frac{1}{2}\mathbf{b}_1, \pm\frac{1}{2}\mathbf{b}_2, ..., \pm\frac{1}{2}\mathbf{b}_N$, where $\mathbf{b}_i$ is a basis vector. Each of these parallelotopes are a translated version of $\frac{1}{2}\mathcal{A}$ with the second moment $\left(\frac{1}{2}\right)^{N+2} X + |\pm\frac{1}{2}\mathbf{b}_1, \pm\frac{1}{2}\mathbf{b}_2, ..., \pm\frac{1}{2}\mathbf{b}_N|^2 \text{volume}(\frac{1}{2}\mathcal{A})$. By the summation over all these second moments, we can find the second moment of $\mathcal{A}$.

$$\begin{aligned} X = \sum_1^{2^N} & \left[\left(\tfrac{1}{2}\right)^{N+2} X + |\pm\tfrac{1}{2}\mathbf{b}_1, \pm\tfrac{1}{2}\mathbf{b}_2, ..., \pm\tfrac{1}{2}\mathbf{b}_N|^2 \text{volume}(\tfrac{1}{2}\mathcal{A})\right] \\ &= \left(\tfrac{1}{2}\right)^2 X + 2^{N-2}(|\mathbf{b}_1|^2 + ... + |\mathbf{b}_1|^2)\text{volume}(\tfrac{1}{2}\mathcal{A}) \\ &= \tfrac{1}{4}X + \tfrac{1}{4}(|\mathbf{b}_1|^2 + ... + |\mathbf{b}_1|^2)\text{volume}(\mathcal{A}) \\ &\Longrightarrow X = \tfrac{1}{3}(|\mathbf{b}_1|^2 + ... + |\mathbf{b}_1|^2)\text{volume}(\mathcal{A}). \end{aligned} \tag{12}$$

References

[1] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiple-antenna multi-user communications-Part II: Perturbation," *Submitted to IEEE Trans. Comm.*, 2003.

[2] G. Caire and S. Shamai, "On the acheivable throughput of a multiple-antenna Gaussian broadcast channel," *IEEE Trans. Info. Theory*, pp. 1691–1706, July 2003.

[3] W. Yu and J. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," in *Proceedings IEEE International Symposium on Information Theory*, p. 498, 2002.

[4] P. Viswanath and D. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Info Theory*, pp. 1912–1921, August 2003.

[5] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, acheivable ratesand sum capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Info. Theory*, pp. 2658–2658, August 2003.

[6] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Info. Theory*, pp. 1250–1276, June 2002.

[7] C. Windpassinger, R. F. H. Fischer, and J. B. Huber, "Lattice-reduction-aided broadcast precoding," in *5th International ITG Conference on Source and Channel Coding (SCC)*, (Erlangen, Germany), pp. 403–408, January 2004.

[8] B. Helfrich, "Algorithms to construct minkowski reduced and hermit reduced lattice bases," *Theoretical Computer Sci. 41*, pp. 125–139, 1985.

[9] A. Korkine and G. Zolotareff, "Sur les formes quadratiques," *Mathematische Annalen*, vol. 6, pp. 366–389, 1873 (in French).

[10] C. P. Schnorr, "A hierarchy of polynomial time lattice basis reduction algorithms," *Theoretical Computer Sci. 53*, pp. 201–224, 1987.

[11] A. K. Lenstra, H. W. Lenstra, and L. Lovasz, "Factoring polynomials with rational coefficients," *Mathematische Annalen*, vol. 261, pp. 515–534, 1982.

[12] M. Taherzadeh, A. Mobasher, and A. K. Khandani, "Communication over MIMO broadcast channels using lattice-basis reduction." Technical Report UW-ECE2004-8.