Partitioned Vector Quantization for MU-MIMO Downlink Broadcasting

Mirza Golam Kibria†, Hidekazu Murata†, Susumu Yoshida†, Koji Yamamoto†, Daisuke Umehara‡
Satoshi Denno*, Masahiro Morikura†

† Graduate School of Informatics, Kyoto University, Kyoto, Japan
‡ Graduate School of Science and Technology, Kyoto Institute of Technology, Japan
* Graduate School of Natural Science and Technology, Okayama University, Japan
Email: †contact-h24e@hanase.kuee.kyoto-u.ac.jp

***Abstract*—A practical and efficient vector quantization called partitioned vector quantization (P-VQ) based non-linear precoder for MIMO downlink broadcasting channels has been analyzed in this paper. P-VQ has been found to be an efficient way of reducing the memory requirements and search complexity in conventional VQ systems, especially for large MIMO. Simulation results reveal that P-VQ technique can enhance the overall quantization performance in terms of achieved capacity, bit error rate (BER) at a cost of reasonable additional complexity under certain feedback budgets. Tomlinson-Harashima precoding (THP), a power and complexity efficient precoding technique along with data dependent vector perturbation (VP) has been employed to meet the power constraint requirements of our system. Least-square (LS) channel estimation is performed at the user terminals to acquire their channel state information (CSI).**

I. Introduction

Multi-user muliple-input multiple-output (MU-MIMO) systems are subjected to a great amount of attention and research due to their inherent capability of enhancing spectral efficiency and link reliability [1]. A transmitter having N_t antennas can serve simultaneously up to N_t users exploiting the same time-frequency resource even when each receiver has only one antenna [2], [3]. Channel state information (CSI) knowledge at base station (BS) offers greater leverage in performance and enhance the benefits of MIMO both for independent and indentically distributed (i.i.d) and non-i.i.d scenarios. Feedback and channel reciprocity are the two fundamental approaches for obtaining CSI at transmitter. Especially in MU-MISO with limited feedback, numerous approaches employ vector quantization techniques to obtain accurate CSI.

The MIMO broadcast system with limited feedback is of great practical interest as it precisely prototypes the realistic systems where individual user feeds back its estimated CSI to the BS. The performance of a vector quantization (VQ) feedback system predominantly depends on how efficiently the quantization vectors span the channel realizations. Better quantization process involves a larger codebook which in turn raises up the feedback load [4] and costs in terms of memory requirement. Again, the quantization error increases proportionally with the number of channel components in each codevector or the number of transmit antenna at BS. For limited feedback based communication systems, feedback budget is a very important factor, as well as the search latency, feedback delay and memory requirements have significant impact on the overall system performance. In this paper, we do performance evaluation of an efficient quantization technique named P-VQ that divides the larger channel matrix into smaller partitions and in turn, reduces the quantization vector exhaustive search complexity, search latency and memory requirements under certain feedback budgets.

In physical design and implementation of a multi-antenna BS, each antenna involves power amplifier in its analog front-end and power handling capacity of this amplifier is exclusively limited due to linear characteristics. Therefeore, power constraint requirements enforced on per-antenna basis is very practical concern in communication systems. A data dependent vector perturbation (VP) precoding [5] combined with non-linear Tomlinson-Harashima precoding (THP) [6] called THP-VP technique has gained a lot of attention as it possesses a structural ability of meeting this transmit power constraint requirements. Though it induces reasonable additional complexity to the BS station, but the receiver complexity is reduced considerably. THP overcomes the drawbacks encountered at receiver side such as noise enhancement, error propagation and power boosting effect at transmitter side.

The remainder of this paper is structured as follows. We discuss the MIMO downlink broadcast system model in section II. Then introduction to the precoding techniques such as THP and VP is provided in section III. Random vector quantization (RVQ) and our proposed P-VQ scheme are discussed in detail in section IV with its structure and signal to interference-plus-noise ratio (SINR) calculation. Section V presents the simulation results and performance analysis followed by some concluding remarks in section VI[1].

II. System Model

We consider a single isolated cell MIMO downlink Gaussian broadcasting communication system with a BS equipped with

[1]**Notation:** Throughout the paper, the operators $(.)^\mathrm{H}$, $E\{.\}$, $(.)^*$ and $(.)^\mathrm{T}$ stand for Hermitian transpose, statistical expectation, conjugate and transpose operation respectively. $\lfloor . \rceil$ refers to rounding to nearest integer of the argument operation. $\mathcal{CN}(\mu, \sigma^2)$ represents the Gaussian distribution of a complex random variable with mean μ and variance σ^2.

978-1-4673-1881-5/12/$31.00

N_t transmit antennas supporting data traffic to $N_r (= N_t)$ non-cooperative and decentralized single-antenna receivers. This broadcasting system is a generalization of MU-MIMO with single-antenna users. For simplifying the performance analysis, block Rayleigh fading frequency flat wireless channel is considered where the channel fading characteristics do not vary within a block but vary randomly from one block to other. The channel output at any time instant t can be written as

$$\boldsymbol{y}(t) = \boldsymbol{H}(t)\boldsymbol{s}(t) + \boldsymbol{v}(t), \tag{1}$$

here, $\boldsymbol{y}(t)$ is $N_t \times 1$ received complex data vector, $\boldsymbol{H}$ is the MU-MISO channel matrix of size $N_r \times N_t$ consisting of complex coefficients considered to be i.i.d and distributed as $\mathcal{CN}(0, 1)$. $\boldsymbol{s}(t) = [s_1, s_2, ..., s_{N_t}]^T$ is the transmitted vector of complex symbols, s_i being the broadcasted symbol intended for user i, generated from memoryless sources and $\boldsymbol{v}(t)$ is length N_r additive white Gaussian noise (AWGN) vector whose components are mutually uncorrelated and have the distribution $\mathcal{CN}(0, \sigma_v^2)$.

We have considered the system to be subjected to sum-power constraint P_T and per-antenna power constraint P_{antenna}, which are governed by these following equations,

$$\begin{aligned} &E[\text{trace}(\boldsymbol{s}(t)\boldsymbol{s}(t)^H)] \le P_T \\ &E[|s_i|^2] \le P_{\text{antenna}} \qquad i = 1.....N_t. \end{aligned} \tag{2}$$

The TH precoding, along with VP becomes very effective when a system has such power constraints and provides significant performance gain [5], [6].

III. Precoding Techniques

In this section, we discuss the structure and operation of THP-VP precoder and formation of its building blocks such as feedback and feedforward filters and the calculation of vector perturbation quantity.

A. Vector Perturbation

VP, a data dependent pre-equalization technique has been found to be superior over the conventional linear pre-equalization techniques and provides significant performance gain. When data dependent VP is employed complying with conventional zero-forcing (ZF) or minimum mean square error (MMSE) pre-equalization criteria, $\boldsymbol{P}_{\text{ZF/MMSE}}$ being the precoding matrix, transmit power is minimized. The vector perturbed transmitted data is expressed as

$$\boldsymbol{s}_{\text{VP}} = \boldsymbol{s} + \tau\zeta, \tag{3}$$

where $\zeta = z + jd$ is a complex number whose components z and d are integers and $\tau = 2(|\alpha|_{\max} + \Psi/2)$, depends on the quadrature amplitude modulation (QAM) constellation scheme. Here $|\alpha|_{\max}$ is the largest constellation point and Ψ is the spacing among the constellation points [5]. The quantity ζ is chosen such that

$$\zeta_{\text{ZF/MMSE}} = \arg\min_{\zeta^*} ||\boldsymbol{P}_{\text{ZF/MMSE}}(\boldsymbol{s} + \tau\zeta^*)||^2. \tag{4}$$

The search for the optimum perturbation quantity ζ is done using lattice reduction based Lenstra-Lenstra-Lovasz (LLL) algorithm which exploits Gram-Schmidt orthogonalization in a repetitive manner [5], [7]. Given a matrix $\boldsymbol{P}_{\text{ZF/MMSE}}$, the LLL algorithm provides $\tilde{\boldsymbol{P}}_{\text{ZF/MMSE}} = \boldsymbol{P}_{\text{ZF/MMSE}}\boldsymbol{T}$, where $\tilde{\boldsymbol{P}}_{\text{ZF/MMSE}}$ is the transformed matrix that has reduced basis, well conditioned. $\boldsymbol{T}$ is an uni-modular matrix with (absolute(determinant($\boldsymbol{T}$)) = 1) that defines the transformation. This transformation does not change or modify the properties of the lattice or original precoding matrix $\boldsymbol{P}_{\text{ZF/MMSE}}$. Again by making use of Babai's approximation [7], the optimum perturbation quantity is obtained as

$$\zeta = \boldsymbol{T}\lfloor \frac{\boldsymbol{T}^{-1}\boldsymbol{s}}{\tau} \rceil. \tag{5}$$

So, after adding the perturbation value $\tau\zeta$ to the original data vector, THP operations are performed as usual.

B. Tomlinson-Harashima Precoding

THP consists of a feedback filter $\boldsymbol{B}$ that pre-distorts the transmitted data eliminating the known causal interference from other users, a modulo operator which keeps the transmitted power within defined range by sending modulo congruent constellation points and a feed forward filter $\boldsymbol{F}$ that confirms causality of the system and de correlate the symbols for different users. QR decomposition of the channel matrix $\boldsymbol{H} = \boldsymbol{RQ}$ produces unit diagonal lower left feedback filter $\boldsymbol{B} = \boldsymbol{GR}$ and unitary feedforward filter $\boldsymbol{F} = \boldsymbol{Q}^H$ where unit gain matrix $\boldsymbol{G}$ is given by $\boldsymbol{G} = \text{diag}(r_{11}^{-1} \cdots r_{kk}^{-1})$ [6], r_{ij} being the $(i, j)^{\text{th}}$ element of matrix $\boldsymbol{R}$.

In a multi-antenna system, the received symbol at each receiver at any time instant t is a superposition of scaled version of all transmitted symbols. The received symbol of user k can be written as

$$y_k = h_{kk}x_k + \sum_{l=1, l\neq k}^{K} h_{kl}x_l + v_k, \tag{6}$$

here, h_{ij} refers to $(i, j)^{\text{th}}$ element in channel matrix $\boldsymbol{H}$. The first term is intended users' symbol and the second term is the accumulated interference from other users.

The non-linearity of TH precoder evolves from the operation of the modulo operator. The cancellation of known multi-user interference is done by the feedback filter $\boldsymbol{B}-\boldsymbol{I}$ [6] performing successive interference cancellation (SIC) as follows

$$x_k = s_{\text{VP}}^k - \sum_{l=1}^{k-1} b_{kl}x_l \qquad k = 1....N_r, \tag{7}$$

here, b_{ij} is the $(i, j)^{\text{th}}$ element of feedback filter. One can comprehend that such pre-cancellation would boost the transmit power, resulting in poor power efficiency. The non-linear modulo operator plays a very important role in keeping the transmit power within defined range. The joint operation of the feedback filter and modulo operator can be shown mathematically as [6]

$$x_k = s_{\text{VP}}^k + x_{\text{THP}}^k - \sum_{l=1}^{k-1} b_{kl}x_l \qquad k = 1...N_r, \tag{8}$$

The modulo operator operates on x_k and adds integer multiples of $x_{\mathrm{THP}} = \pm 2\sqrt{M}$ to confine the transmit power depending on the constellation scheme used. Here, M refers to the order of QAM constellation. After SIC, the filtered data is further passed through the feedforward filter $\boldsymbol{F}$ and gain matrix $\boldsymbol{G}$.

IV. Vector Quantization

Vector quantization is a very practical and efficient way of acquiring downlink CSI, especially in downlink broadcasting scenarios. In our considered finite (the number of receivers is less or equal to the number of transmitting antenna, no multiuser diversity) MU-MISO, the system performance is intensely susceptible or sensitive to the exactness of CSI. Hence, it requires high-rate feedback as the deficient CSI causes multi-user interference [4] which cannot be resolved at the receiver side having only one antenna. And still multi-user interference is not completely removed due to the presence of quantization error, as the quantized channels are no more the actual channels.

Since the generation of the optimal codebook is very difficult, RVQ has been employed, as random quantization codebooks can be generated easily from statistical channel model parameters. Furthermore, RVQ process can simulated efficiently making use of the quantization error statistics when the size of the codebook is very large [4]. In RVQ, the selection of each quantization vector is done independently from N_{t} dimensional unit sphere $C^{N_{\mathrm{r}}\times 1}$, having isotropic distribution. The quantization vectors of different users codebook are independent. Under these circumstances, generation of codebooks are done by providing random rotation to a basis codebook. The MIMO channel matrix $\boldsymbol{H}$ can be expressed as,

$$\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, ..., \boldsymbol{h}_{N_{\mathrm{r}}}]^{\mathrm{T}}, \tag{9}$$

here h_k represents k^{th} user channel vector consisting of N_{t} elements. The codebook $\mathcal{F}$ of size $N_{\mathrm{t}} \times 2^R$ is given by

$$\mathcal{F} = [\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3, ..., \boldsymbol{c}_{2^R}], \tag{10}$$

where $\boldsymbol{c}_i$ is a code vector of size same as user channel vector and R is the number of bits for indexing the code vectors.

Individual user uses separate and independent codebook to prevent the situation when two or more users quantize their respective channels to the same quantization vector, thus limiting the spatial diversity gain. Each user maps its channel vector to the quantization codebook based on minimum distance criterion [3]

$$\hat{\boldsymbol{h}}_k = \boldsymbol{c}_n, \qquad n = \arg \max_{i=1...2^R} |\tilde{\boldsymbol{h}}_k \boldsymbol{c}_i^{\mathrm{H}}|, \tag{11}$$

where $\tilde{\boldsymbol{h}}_k$ is defined by $\boldsymbol{h}_k/||\boldsymbol{h}_k||$, also known as channel direction information (CDI). CSI can be decomposed into two measures namely channel quality information (CQI) and CDI. Hence the complete CSI is expressed as $CSI_k = \tilde{h}_k.||\hat{h}_k||$ where $||\hat{h}_k||$ is the norm based CQI. There has been a lot of research work to design efficient codebooks and some fundamental characteristic features are also established. Those characteristic features and results motivate us to employ P-VQ in our system for reducing the codebook size and search latency. The two main features of vector quantization are (i) there is a linear growth in feedback requirements with increasing SNR [8] and (ii) codebook size increases drastically as N_{t} increases and some results are derived in [9]

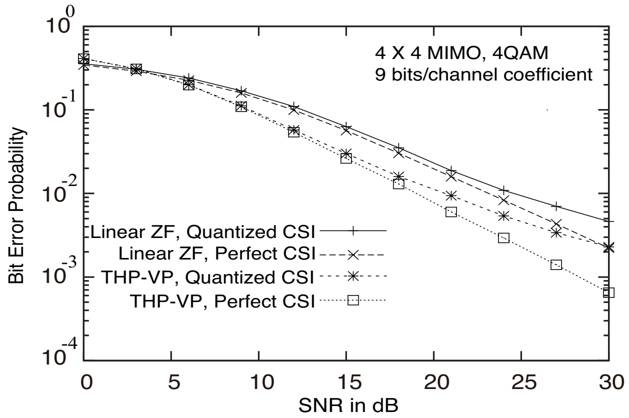

Fig. 1. BER performance comparison for linear and non-linear precoding techniques.

A. Partitioned Vector Quantization P-VQ

Considering all the facts discussed in the previous sections, we thus propose the use of a simple, practical and yet very efficient quantization process named P-VQ [10], where users channel vectors are partitioned into two or more blocks (yet not violating the definition of a vector) and then quantized independently. This method is very efficient both for large and moderate size MIMO systems. As a result, the codebook size drops down by a large factor and search latency, as well as complexity reduces proportionally[2]. The channel matrix $\boldsymbol{H}$ can also be expressed as

$$\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{h}_2, ..., \boldsymbol{h}_{N_{\mathrm{t}}}], \tag{12}$$

where $\boldsymbol{h}_n$ is a $N_{\mathrm{r}} \times 1$ size column vector. These N_{t} column vectors are partitioned into $N(\geq 2)$ different blocks of size $N_{\mathrm{r}} \times (N_{\mathrm{t}}/N)$ each if N_{t} is even. All blocks may or may not be of same size (depends on the divisibility of N_t by N). As a result, in P-VQ, instead of having only one channel vector per user, each user has N channel vectors and each channel vector is quantized independently as in conventional RVQ. For example, when $N_{\mathrm{t}} = 8$ and $N = 4$, channel matrix $\boldsymbol{H}$ (from (12)) can be partitioned as

$$\boldsymbol{H} = [\boldsymbol{h_1}, \boldsymbol{h_2}|\boldsymbol{h_3}, \boldsymbol{h_4}|\boldsymbol{h_5}, \boldsymbol{h_6}|\boldsymbol{h_7}, \boldsymbol{h_8}]. \tag{13}$$

[2]We base our P-VQ assumptions on the fact that when the channel coefficients i.i.d and R bits indexing (codebook size 2^R) is required to find a mapped quantization value with small distortion for one channel element (SISO channel), then for a MISO channel with N_{t} i.i.d elements (also known as downlink vector channel), the codebook size becomes $2^{RN_{\mathrm{t}}}$. Whereas, after partitioning N_{t} into N equal partitions, each partition containing $\frac{N_{\mathrm{t}}}{N}$ elements, the codebook size for one partition becomes $2^{\frac{N_{\mathrm{t}}}{N}R}$. So,the combined codebook size for partitioned case is $N(2^{\frac{N_{\mathrm{t}}}{N}R}) \leq 2^{RN_{\mathrm{t}}}$, where $N \geq 2$.

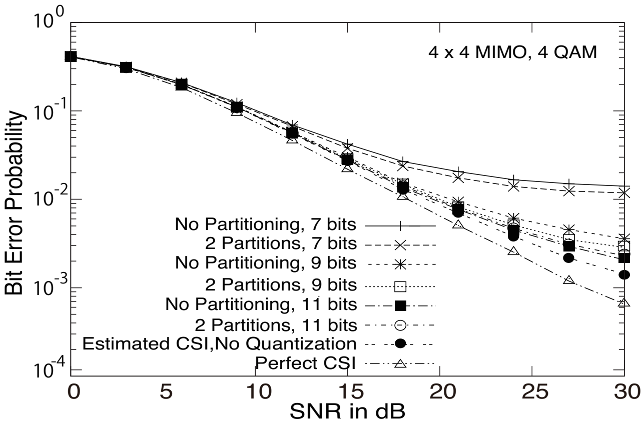

Fig. 2. BER performance comparison for different feedback budgets with and without partitioning.

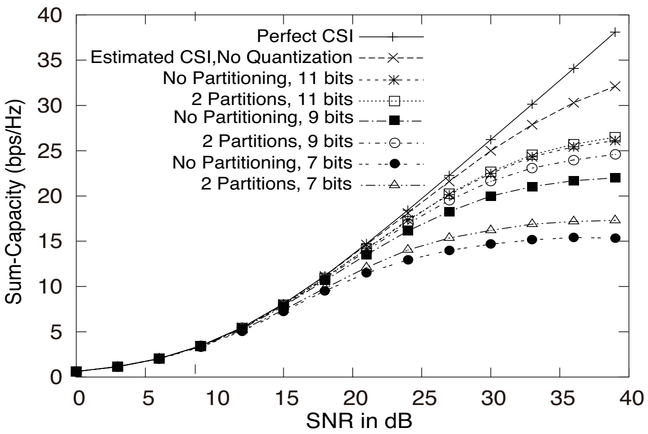

Fig. 3. Sum-capacity comparison for different feedback budgets with and without partitioning for 4×4 MIMO.

which is analogous to the case where each user contains N independent small sized codebooks

$$\mathcal{F}_k = [\mathcal{F}^1{}_k|\mathcal{F}^2{}_k|...|\mathcal{F}^N{}_k]. \tag{14}$$

instead of having a single very large codebook. As CQI distribution (mean and variance) does not change for small size partitions, we reserve a fixed amount of bits for CQI quantization both in conventional VQ and our proposed P-VQ. And also for fair comparison, we consider the same feedback budget for both cases. The total feedback budget $\mathcal{B}$ bits are distributed among all the partitions as follows,

Total feedback budget = $\mathcal{B}$

Feedback budget per user, $\mathcal{B}_{\text{user}} = \mathcal{B}/N_{\text{r}}$

CQI quantization budget = $\mathcal{B}_{\text{CQI}}$

Number of partitions = N

VQ CDI budget/channel coefficient = $(\mathcal{B}_{\text{user}} - \mathcal{B}_{\text{CQI}})/N_{\text{t}}$

P-VQ CDI budget/channel coefficient = $(\mathcal{B}_{\text{user}} - \mathcal{B}_{\text{CQI}}N)/N_{\text{t}}$

After the quantization vector indexes are obtained for all the partitions, they are put in order and sent back to BS.

B. SINR Calculation and Sum-rate Analysis

The transmitted data first passes through the feedback filter, where SIC is performed to remove the known multi-user interference. Then it passes through the feedforward filter, defined as $\boldsymbol{F} = \boldsymbol{Q}^{\text{H}}$ and AWGN noise is added. So, the received data at the receiver can be expressed as.

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{Q}^{\text{H}}\boldsymbol{s} + \boldsymbol{v}. \tag{15}$$

We mention here that the channel that the user quantizes is not the original channel, but the LS estimated channel. So, the sum capacity is affected not only by the quantization error, but also channel estimation error which is the most probable case in practical applications.

$$\begin{aligned}\boldsymbol{y} &= (\hat{\boldsymbol{H}}_{\text{est}} + \Delta_{\text{est}})\boldsymbol{Q}^{\text{H}}\boldsymbol{s} + \boldsymbol{v},\\ &= ((\hat{\boldsymbol{H}}_{\text{quant}} + \Delta_{\text{quant}}) + \Delta_{\text{est}})\boldsymbol{Q}^{\text{H}}\boldsymbol{s} + \boldsymbol{v},\\ &= (\hat{\boldsymbol{H}}_{\text{quant}} + \Delta_{\text{total}})\boldsymbol{Q}^{\text{H}}\boldsymbol{s} + \boldsymbol{v},\\ &= \boldsymbol{R}\boldsymbol{s} + \Delta_{\text{total}}\boldsymbol{Q}^{\text{H}}\boldsymbol{s} + \boldsymbol{v}.\end{aligned} \tag{16}$$

here $\Delta_{process}$ refers to the error induced by *process*. And, the data received by individual user can be expressed as

$$y_k = r_{kk}s_k + \sum_{j\neq k} r_{kj}s_j + (\Delta^k_{\text{total}})^{\text{T}}\boldsymbol{Q}^{\text{H}}s + v_k, \tag{17}$$

The second part is the multi-user interference which is already taken care of by SIC process. The feed-forward matrix $\boldsymbol{Q}^{\text{H}}$ is an unitary matrix, hence no amplification of the third term in above expression.

So, the signal-to-interference-ratio (SINR) per user is written as

$$SINR_k = \frac{r_{kk}^2 P_k}{N_0 + \sum_{j\neq k}||\Delta^j_{\text{total}}||^2 P_j}, \tag{18}$$

Finally, the sum-rate is given by

$$C_{\text{sum-rate}} = \sum_{k=1}^{N_{\text{r}}} E\left\{log_2\left(1 + \frac{r_{kk}^2 P_k}{N_0 + \sum_{j\neq k}||\Delta^j_{\text{total}}||^2 P_j}\right)\right\}. \tag{19}$$

V. Simulation Parameters and Results

We have considered two different sized MIMO channel matrices as 4×4 and 8×8 for analysis and 4 QAM has been used as the modulation scheme for all simulations. We consider the number of partitions to be 2 for 4×4 MIMO and 2, as well as 4 for 8×8 MIMO. Block length of 1000 symbols is used. 6 bits are always reserved for norm based CQI quantization. To restrict our system from becoming interference limited, a good estimate of CSI is achieved by quantizing the channels using a rich quantization codebook. The RVQ process for large size codebook is simulated exploiting the quantization error statistics as described in [4].

Fig. 1 depicts BER performance comparison between linear and non-linear precoding schemes. It shows that non-linear precoding clearly has an edge over linear precoding. Though the gain achieved in terms of BER and capacity is not very high, yet the system power constraint requirements are fulfilled. In Fig. 2 and Fig. 3, we observe that our proposed P-VQ has degraded performance in terms of BER and sum-capacity brought by complexity advantages for 4×4 MIMO over conventional VQ, when the feedback budget per channel coefficient is ≤ 6. This is due to the fact that, for an equal feedback budget, the single codebook size for conventional VQ is very large and reasonably sufficient for good quantization, but for P-VQ, the individual codebooks are very small as per calculation given in section V, and thus insufficient for finer quantization. Hence the accumulated quantization error from all the partitions in P-VQ is larger than that of VQ using a single codebook. When the amount of feedback bits is further increased, our P-VQ outperforms conventional VQ. And it is clearly reflected in Figs. 2 and 3. However, if we consider the search latency involved in finding the optimum codevector through exhaustive search both in conventional VQ and P-VQ, the latter would be a suitable choice.

We also figure out in Figs. 4 and 5 that the capacity growth for 8×8 is comparatively slower than that of 4×4 channel matrix for relatively smaller values of feedback bits as the quantization error is severe for larger dimensional MIMO with low feedback budget [9]. Both for 4×4 and 8×8 MIMO, at SNR = 30dB, the optimal value of feedback bits per channel is 10-11bits (approximately), after which the performance difference between partitioned VQ and non-partitioned is not significant. For 8×8 MIMO, after it reaches the optimum capacity when feedback bits/channel coefficient=10 bits, there is a slight reduction in capacity. Clearly, there exists a partitioning zone where switching between VQ and P-VQ can be done depending on feedback budget.

In Fig. 5, we analyze the effect of partition number N on sum-capacity for 8×8 MIMO. For less amount of feedback budget, 2 partitioned curve fits well with the non-partitioned capacity curve compared to 4 partitioned case. It may be due to the fact that, at that stage, the accumulation of quantization errors from all 4 partitions is larger than that of 2 partitions case. Still our P-VQ performs reasonably well if we consider the memory requirements and search complexity. As the amount of feedback bits are increased, the 4 partitioned capacity gets enhanced.

VI. Conclusions

Simulation results show that the P-VQ technique can greatly reduce the quantization vector search complexity with small performance degradation while the feedback budget is small. When the feedback-rate is further increased, P-VQ outperforms conventional VQ in terms of BER and capacity. For large dimensional MIMO, more number of partitions is preferred when the system has comparatively high feedback budget. Above all, P-VQ reduces the codevector search complexity and latency significantly.

Acknowledgment

This work was supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications, Japan.

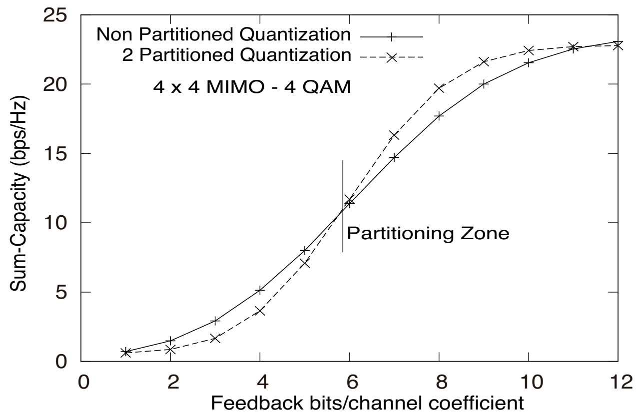

Fig. 4. Sum-capacity comparison between partitioned and non-partitioned quantization with different feedback budgets for 4×4 MIMO.

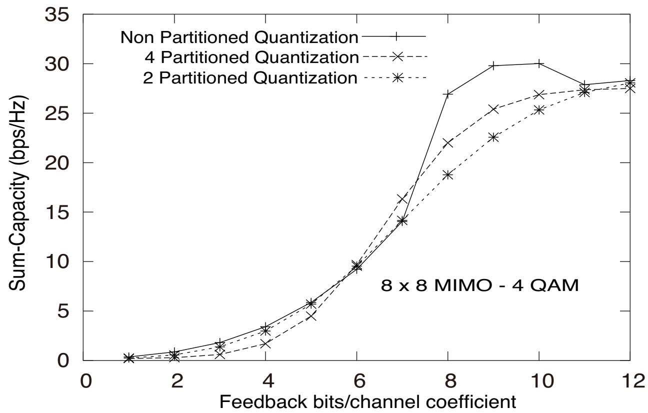

Fig. 5. Sum-capacity comparison between partitioned and non-partitioned quantization with different feedback budgets for 8×8 MIMO.

References

[1] G.J. Foschini and M.J. Gans,"On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun. : Kluwer Academic Press.*, no. 6, pp. 311-335, 1998.

[2] G. Caire and S. Shamai,"On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. on Info. Theory.*,vol. 49, no. 7, pp. 1691-1706, Jul. 2003.

[3] T. Yoo, N. Jindal and A. Goldsmith,"Multi Antenna Downlink Channels with Limited feedback and User Selection," *IEEE Journals on Selected Areas in Communication*, vol. 25, no. 7, pp. 1478-1491, Sep. 2007.

[4] N. Jindal,"Antenna Combining for the MIMO Downlink Channel," *IEEE Trans. on Wireless Commun.*, vol. 7, no. 10, pp. 3834-3844, Oct. 2008.

[5] M. Airy, S. Bhadra,R. W. J. Heath and S. Shakkattai,"Transmit Precoding for the Multiple Antenna Broadcast Channel," *IEEE 63rd Vehicular Technology Conference-Melbourne*,pp. 1396-1400,May. 2006.

[6] R. F. H. Fischer, C. Windpassinger,A. Lampe and J. B. Hubber, "Space-Time Transmission using Tomnlinson-Harashima Precoding," *Chair of Information Transmission,University of Erlangen-Nurnberg* , Germany.

[7] A.B. Gershman and N.D. Sidiropoulos,"Space-time Processing for MIMO Communications," *John Wiley and Sons Limited.*

[8] L. Xinmin and B. Baoming,"ZF-THP Combined with Receive Beamforming for Multi-user MIMO Downlinks," *IEEE ICISE, Xian, China*, vol. 7, no. 10, pp. 2779-2782, Dec. 2009.

[9] S. Zhou, Z. Wang and G.B. Giannakis,"Performance Analysis for Transmit-Beamforming with Finite-Rate Feedback," *Conference on Information Science and Systems, Princeton*, pp. 880-885, Mar. 2009.

[10] G. Motta, F. Rizzo and J.A. Storer,"Partitioned Vector Quantization: Application to Lossless Compression of Hyperspectral Images,"*IEEE ICASSP Proceedings.*, vol. 3, pp. III-241-244, Apr. 2003.