COMPLEXITY REDUCTION FOR VECTOR PRECODING USING QOS REQUIREMENTS

Christos Masouros[1], *Tharm Ratnarajah*[2], *Mathini Sellathurai*[3]

[1]Dept. of Electrical & Electronic Eng., University College London, London, UK
[2]School of Engineering, University of Edinburgh, Edinburgh, UK
[3]School of Eng. and Physical Sciences, Herriot Watt University, Edinburgh, UK

ABSTRACT

We propose a low-complexity vector precoding (VP) scheme for the downlink of multi-user multiple input multiple output (MU-MIMO) systems. Instead of performing a full sphere search to maximize the receive signal to noise ratio (SNR), the search for the perturbation vectors finishes once a threshold SNR value is reached, thus saving significant computational burden at the transmitter. This threshold is determined by the quality of service (QoS) requirements of the mobile users. To evaluate the advantages of the proposed technique compared to VP, we analytically calculate its computational complexity in terms of the volume of the associated search space. The results show that the proposed thresholded VP (TVP) offers a significantly reduced complexity compared to VP.

***Index Terms*—** Vector precoding, sphere encoding, complexity reduction, multi-user MIMO, non-linear precoding

1. INTRODUCTION

The pursuit of cost- and power-efficient mobile units has stimulated a growing interest in precoding schemes for the downlink MU-MIMO transmission. Capacity achieving non-linear dirty paper coding (DPC) techniques [1, 2] are in general impractical as they assume codewords with infinite length for the encoding of the data. Their suboptimal counterparts [3, 4] offer a complexity reduction at a comparable performance. Still however, the associated complexity is prohibitive for their deployment in current communication standards. On the other hand, linear precoding schemes based on channel inversion [5, 6, 7] offer the least complexity, but poor performance.

A performance improvement is provided by vector precoding (VP) [8], by judiciously perturbing the data vectors at the transmitter. This results in much enhanced receive signal to noise ratios (SNRs) compared to linear precoding. The improved performance however, comes at the expense of an increased complexity since the search for the optimal perturbation vectors is an NP-hard problem, typically solved by sphere search algorithms at the transmitter.

*This work was supported by the Royal Academy of Engineering, UK and EPSRC grant EP/I037156/1.

The complexity of various sphere search techniques has been studied in [9]-[10] (among others) in terms of search nodes visited and search lattice volumes. A number of techniques have been proposed towards reducing the complexity of VP precoding (e.g. [11]-[13]). In [11] a search over a reduced lattice is proposed, based on empirical observations of the relation between the instantaneous symbols and the optimum perturbation vectors. Further work in [12] has proposed the decoupling of the perturbation optimization in the real and imaginary domain of the data symbols thus offering a lower complexity compared to the joint optimization approach.

In this paper we propose a thresholded approach, where we apply a threshold to the VP optimization and a resulting search-termination threshold to the sphere encoder. The aim is to reduce the associated precoding complexity by reducing the number of nodes that need to be visited by the sphere encoder, for a given target performance. The threshold value is directly determined by the operational SNR requirement of the mobile users. To do this we adapt the Schnorr-Euchner search (SE) algorithm [10] by introducing a threshold on the required weight (precoded signal norm) of the optimal node. If this threshold weight is met, the search is terminated at a reduced number of nodes visited.

It should be noted that the proposed can be applied on top of other complexity reduction techniques to further reduce complexity by means of a thresholded search, compared to the case where the full sphere search is carried out. However, to keep the focus of this work on the central idea we only use conventional VP [8] as the reference for comparison.

2. SYSTEM MODEL AND VECTOR PRECODING

Let us assume a downlink system with a single BS equipped with N transmit antennas and $M \leq N$ single-antenna users. The received signal is given as

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{x} + \mathbf{w} \tag{1}$$

where $\mathbf{r} \in \mathbb{C}^{M \times 1}$ and $\mathbf{H} \in \mathbb{C}^{M \times N}$ is the channel matrix. Also, $\mathbf{x} \in \mathbb{C}^{N \times 1}$ is the transmit symbol vector $\mathbf{w} \in \mathbb{C}^{M \times 1} \sim \mathcal{CN}(0, \sigma^2 \mathbf{I}_M)$ is the vector of the additive white Gaussian noise (AWGN).

978-1-4799-0356-6/13/$31.00 ©2013 IEEE

VP precoding employs a channel inversion precoding matrix and applies a perturbation on the transmitted symbols such that the signal content at the receiver is maximized. The transmitted signal is given by [8]

$$\mathbf{x} = \sqrt{\frac{P}{\beta}}\mathbf{H}^{\dagger}(\mathbf{u} + \tau\mathbf{l}^{\star}) \tag{2}$$

where $\mathbf{H}^{\dagger}$ is the Moore-Penrose generalized inverse of matrix $\mathbf{H}$, $\mathbf{u} \in \mathbb{C}^{M\times 1}$ is the data symbol vector,

$$\beta = ||\mathbf{H}^{\dagger}(\mathbf{u} + \tau\mathbf{l}^{\star})||^2 \tag{3}$$

is the transmit power scaling factor so that $E\{||\mathbf{x}||^2\} = P$ and $\mathbf{l}^{\star} \in \mathbb{C}^{M\times 1}$ is the selected perturbation vector with integer entries. Also $\tau = 2|c|_{max} + \Delta$ where $|c|_{max}$ is the absolute value of the constellation symbol with the maximum magnitude and Δ denotes the minimum Euclidean distance between constellation symbols. To maximize the signal component in the received symbols or equivalently minimize the noise amplification, the perturbation vectors $\mathbf{l}^{\star}$ are chosen as

$$\mathbf{l}^{\star} = \arg\min_{\mathbf{l}\in\mathbb{Z}^M+\mathrm{j}\mathbb{Z}^M} ||\mathbf{H}^{\dagger}(\mathbf{u} + \tau\mathbf{l})||^2 \tag{4}$$

For complex symbol alphabets, the optimization in (4) is a $2M$-dimentional real integer lattice problem, known to be NP-hard. Sphere search techniques are typically employed to solve the minimization, the complexity of which is known to grow exponentially with M. Based on the above expressions the received symbol vector can be calculated as

$$\mathbf{r} = \sqrt{\frac{P}{\beta}}(\mathbf{u} + \tau\mathbf{l}^{\star}) + \mathbf{w} \tag{5}$$

At the receiver, the signal is first scaled back to eliminate the effect of the transmit scaling factor and then fed to a modulo operator to remove the perturbation quantity $\tau\mathbf{l}^{\star}$. The output of the modulo stage is given as

$$\mathbf{y} = \mathrm{mod}_{\tau}\left[\sqrt{\frac{\beta}{P}}\mathbf{r}\right] = \mathrm{mod}_{\tau}\left[\mathbf{u} + \tau\mathbf{l}^{\star} + \sqrt{\frac{\beta}{P}}\mathbf{w}\right] = \mathbf{u} + \mathbf{n} \tag{6}$$

where

$$\mathrm{mod}_{\tau}(x) = x - \left\lfloor\frac{\Re(x) + \tau/2}{\tau}\right\rfloor\tau - j\left\lfloor\frac{\Im(x) + \tau/2}{\tau}\right\rfloor\tau \tag{7}$$

is the modulo operation with base τ, $\Re(x)$ and $\Im(x)$ denote the real and imaginary parts of x respectively. Also, $\lfloor x \rfloor$ denotes the maximum integer less or equal to x and vector $\mathbf{n}$ in (6) denotes the equivalent noise vector at the receiver after the scaling and modulo operation.

3. PROPOSED LOW COMPLEXITY VECTOR PRECODING

Typically, the precoder is operating under a threshold performance requirement from the mobile users. Accordingly, assume that the received SNR γ excluding the modulo operation is required to be greater or equal to a certain threshold γ_t

$$\gamma = \frac{E_b}{N_0}\cdot\frac{P}{\beta} \geq \gamma_t \tag{8}$$

where E_b, N_0 are the per bit power and noise power respectively. Denoting the complexity of the precoder as C, we formulate the optimization problem as

$$\begin{aligned}&\min C(\beta_t)\\ &\text{s.t.c. } \beta(\mathbf{l}^{\star}) \leq \frac{E_b}{N_0}\cdot\frac{P}{\gamma_t} \triangleq \beta_t\end{aligned} \tag{9}$$

where we have introduced a signal norm threshold β_t based on the SNR requirement of (8). Hence, the above optimization minimizes the precdoing complexity, subject to a receive SNR threshold. Since there is no exact closed form expression of the complexity associated with VP, the optimization cannot be solved directly. To solve the optimization, we adapt the sphere encoder such that the search is terminated once the first perturbation vector that is found to satisfy the weight threshold β_t. In the case when the instantaneous channel is such that the threshold can not be met, the perturbation corresponding to the minimum weight is selected according to the conventional sphere encoding process. The selected perturbation vector is therefore given as

$$\mathbf{l}^{\star} = \begin{cases}\mathbf{l}_t, \text{if } \exists\mathbf{l}_t : ||\mathbf{H}^{\dagger}(\mathbf{u} + \tau\mathbf{l}_t)||^2 \leq \beta_t\\ \arg\min_{\mathbf{l}\in\mathbb{Z}^M+\mathrm{j}\mathbb{Z}^M} ||\mathbf{H}^{\dagger}(\mathbf{u} + \tau\mathbf{l})||^2, \text{otherwise}\end{cases} \tag{10}$$

Apart from the above modification on the sphere search, the transmit and receive processing for TVP is identical to the one of VP given in (2)-(7).

4. COMPLEXITY ANALYSIS

It has been observed in [10] and references therein that the complexity of the SE search is proportional to the volume of the region being searched in the lattice space. At the k-th search layer of the tree search this volume is a hypersphere with maximum radius α_k. A geometrical illustration of this is shown in fig. 1(a) for a 2-dimensional lattice. Here the dots represent the lattice points and the area inside the outer circle denotes the area of candidate points searched, based on the search radius.

For the proposed thresholded SE (TSE), due to the applied threshold, the search is carried out to the point where β_t is satisfied and no further lattice points closer to the target point need to be evaluated. The equivalent search space for the 2-dimentional lattice is denoted by the shaded area in fig. 1(b).

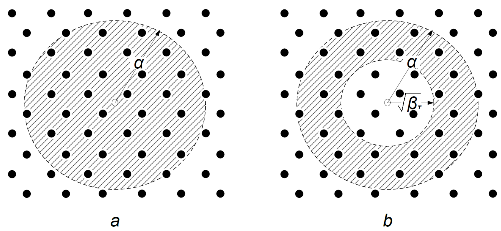

Fig. 1. Geometrical representation of the 2-dimensional integer-lattice sphere search problem for a) conventional sphere search and b) thresholded sphere search

It is therefore clear that, the search volume associated with TSE is the volume of the hypersphere with radius α_k less the volume of the hypersphere with radius $\sqrt{\beta_t}$. When the termination threshold β_t is too strict and a full SE search is carried out, this signifies that the equivalent volume of the threshold hypersphere is zero, as it contains zero lattice points. Note that, since the TSE terminates once the *first* lattice point satisfying the threshold is found, the area shown in fig. 1(b) consists of an upper bound of the actual volume searched for TSE. Typically TSE will only search a part of this volume until the first lattice point on the surface of the inner hypersphere is met.

In the following, we use these key observations to attain a complexity evaluation for TSE, following the methodology of [9, 10]. We use $r_{n,m}$ to denote the n, m-th element of matrix $\mathbf{R}$ of the QR decomposition $QR(-\mathbf{H}^{\dagger}\tau)$.

Theorem 1 *Denote the search volume for the k-th search layer with radius α_k as $\mathcal{V}_k(\alpha_k)$. For the complexity $C(\beta_t)$ associated with TSE in a M-dimensional lattice we have*

$$C(\beta_t) \leq C \propto V = \left[\mathcal{V}_M(\infty) - \mathcal{V}_M(\sqrt{\beta_t})\right]^+ \tag{11}$$

for which

$$\left[\mathcal{V}_M(\infty) - \mathcal{V}_M(\sqrt{\beta_t})\right]^+ \leq \frac{\pi^{M/2}\left[\left(\frac{\phi_M}{2}\right)^M - \beta_t^{M/2}\right]^+}{\Gamma(M/2+1)} \tag{12}$$

where $\phi_k = \sqrt{r_{1,1}^2 + r_{2,2}^2 + \cdots + r_{k,k}^2}, k = 1, \ldots, M$, $[x]^+ = \max\{0, x\}$ *and* $\Gamma(.)$ *denotes the gamma function.*

Proof: Firstly, (11) is a direct expression of the fact that, for TSE, according to fig. 1(b) and the above discussion, the complexity upper bound C is proportional to the volume of the hypersphere with radius α_k less the volume of the hypersphere with radius $\sqrt{\beta_t}$. If $\sqrt{\beta_t} > \alpha_k$ (i.e. the $\sqrt{\beta_t}$-hypersphere contains the α_k-hypersphere) then the search volume is zero: all lattice points within the hypersphere with radius α_k satisfy the weight threshold and a random lattice point can be chosen.

It is shown in [10] that the volume of the hypersphere with radius α_k follows

$$\mathcal{V}_k(\alpha_k) \leq \frac{\pi^{k/2}}{\Gamma(k/2+1)}\prod_{n=1}^{k}\alpha_n \leq \frac{\pi^{k/2}}{\Gamma(k/2+1)}\alpha_k^k \tag{13}$$

which also corresponds to the volume of the search space for conventional VP. The second inequality in (13) stems from the fact that the volume of a hypersphere with radius α_n in dimension $n = 1, \ldots k$ is less than the volume of the hypersphere with constant radius $\alpha_k \geq \alpha_n, \forall n = 1, \ldots k$ in all dimensions. Since the initial search radius for the SE algorithm is unbounded, the search volume is upper-bounded by $\mathcal{V}_k(\infty)$. Using (13) and lattice theory we have the following upper-bound for the first term on the right side of (11)

$$\mathcal{V}_k(\infty) \leq \frac{\pi^{k/2}}{\Gamma(k/2+1)}\prod_{n=1}^{k}\frac{\phi_n}{2} \leq \frac{\pi^{k/2}}{2^k\Gamma(k/2+1)}\phi_k^k \tag{14}$$

where ϕ_k is defined as in the theorem above. For the inner hypersphere of fig. 1(b) with radius $\tilde{\alpha}_1 = \sqrt{\beta_t}$ at the leaf level of the sphere search we have the following recursion

$$\mathcal{V}_k(\sqrt{\beta_t}) = \int_{-\tilde{\alpha}_k}^{\tilde{\alpha}_k}\mathcal{V}_{k-1}\left(\sqrt{\tilde{\alpha}_k^2 - y^2}\right)dy, k = 2, \ldots, M \tag{15}$$

where $\tilde{\alpha}_k$ is the radius for the k-th layer of TSE, with starting volume

$$\mathcal{V}_1(\sqrt{\beta_t}) = 2\sqrt{\beta_t} \tag{16}$$

This recursion yields

$$\mathcal{V}_k(\sqrt{\beta_t}) = \frac{\pi^{k/2}}{\Gamma(k/2+1)}\sqrt{\beta_t}\prod_{n=2}^{k}\tilde{\alpha}_n \geq \frac{\pi^{k/2}}{\Gamma(k/2+1)}\beta_t^{k/2} \tag{17}$$

In (17) the term on the right denotes the volume of a hypersphere with constant radius $\sqrt{\beta_t}$ which is the minimum radius in the TSE recursion. Substituting (17) and (14) for $k = M$ in (11) yields (12) which concludes the proof.

It was shown in [9] that for an infinite lattice the expected number of lattice points contained inside a k-dimensional hypersphere of radius α is given as

$$E_p\left(k, \alpha^2\right) = \sum_{q=0}^{\infty}\varphi\left(\frac{\alpha^2}{2(\sigma^2+q)}, \frac{k}{2}\right)\cdot r_k(q) \tag{18}$$

where $\varphi\left(x, \kappa\right)$ is the normalized incomplete gamma function and $r_\kappa(x)$ denotes the number of ways a non-negative integer x can be represented as a sum of κ squares of integers [9]. Accordingly, it can be shown that the expected number of lattice points visited by the TSE search at the k-th layer is upper bounded by

$$p(k) = \sum_{q=0}^{\infty}\varphi\left(\frac{[\phi_k^2/4 - \beta_t]^+}{2(\sigma^2+q)}, \frac{k}{2}\right)\cdot r_k(q) \tag{19}$$

The complexity in numbers of numerical operations can then be calculated using the formula

$$C(\beta_t) \leq \sum_{k=1}^{M} f_p(k)\sum_{q=0}^{\infty}\varphi\left(\frac{[\phi_k^2/4 - \beta_t]^+}{2(\sigma^2+q)}, \frac{k}{2}\right)\cdot r_k(q) \tag{20}$$

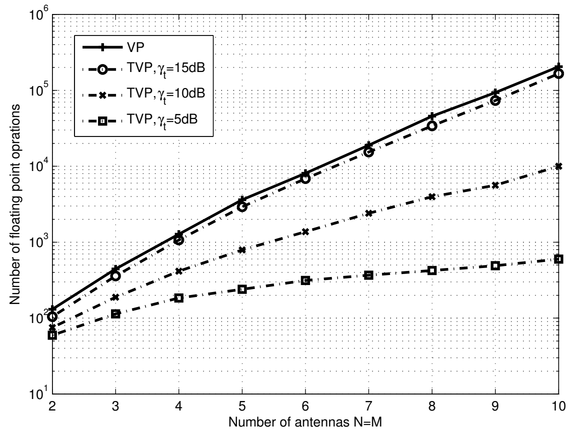

Fig. 2. NOP vs. $N = M$ for VP and TVP with γ_t=5dB, 10dB, 15dB for txSNR=20dB, 4QAM

where $f_p(k) = 2k + 11$ denotes the number of numerical operations per visited node [9] in the k-th search layer.

Since $\mathcal{V}_M(\sqrt{\beta_t}) \geq 0$ and $E_p(k, \beta_t) \geq 0$, it can be concluded that the proposed TVP has a strictly reduced complexity compared to VP.

5. NUMERICAL RESULTS

To evaluate the complexity gains of the proposed scheme, we illustrate Monte Carlo simulations of the proposed TVP and conventional VP for the frequency flat Rayleigh fading statistically uncorrelated MIMO channel whose impulse response is assumed perfectly known at the transmitter. Fig. 2 shows the complexity of VP and TVP in terms of numbers of floating point operations (NOP) for increasing numbers of antennas $N = M$, with $\rho_{dB} = 20dB$. At this point exact NOP counts are shown based on simulation, while the upper bounds derived in section 4 are compared to simulation in following results. The cases of three different receive SNR thresholds are observed $\gamma_t = 5$dB, $\gamma_t = 10$dB, $\gamma_t = 15$dB. It can be seen that for $\gamma_t = 15$dB only a minor complexity reduction is achieved for the cases with well conditioned channel matrices. The complexity benefits however improve as the SNR thresholds are reduced, allowing for a more relaxed precoding optimization in (10).

In fig. 3 the complexity in numbers of operations is shown for increasing values of γ_t in dB. For high performance thresholds, TVP has the same complexity as VP, since full SE search is carried out as the weight threshold cannot be satisfied. Complexity benefits can be observed as γ_t reduces below 20dB. The theoretical complexity upper bounds for VP from [10] and TVP from section 4 are included for the 4×4 case for comparison to the numerical results. The difference between theoretical and simulated complexity, as mentioned above is due to the fact that constant radius hyperspheres with the maximum

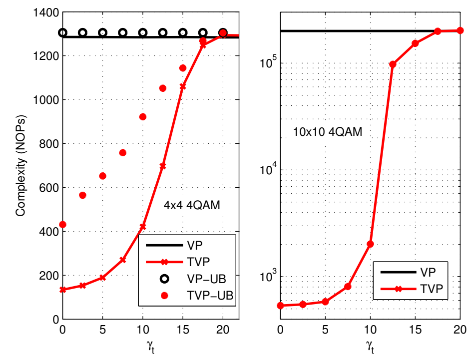

Fig. 3. NOP vs. γ_t for VP and TVP with txSNR=20dB, 4QAM

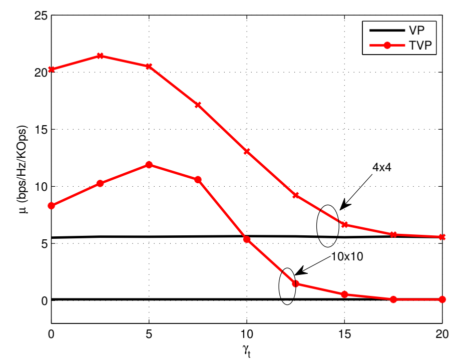

Fig. 4. μ vs. γ_t for VP and TVP with txSNR=20dB, 4QAM

observed radius are considered in the analysis for both VP and TVP (as opposed to adaptive radius used in SE). Furthermore, for TVP the volume and number of lattice points calculated is an overestimate of the ones visited, as the search terminates once the first lattice point on the surface of the threshold hypersphere is met.

Next, fig. 4 shows the tradeoff between complexity and performance. We introduce the relevant metric $\mu = \frac{R}{C}$ in units of bps/Hz per 10^3 operations (KOps), where R is the obtained sum rate and C the associated complexity in numbers of operations. For the 4×4 MIMO the value $\gamma_t = 2.5$dB maximizes the tradeoff metric, while for the 10×10 case the optimum value increases to $\gamma_t = 5$dB.

6. CONCLUSION

We have shown that significant computational gains can be achieved for vector precoding for given user performance requirements. The complexity efficiency of both conventional and proposed techniques was studied in terms of the volume of the search space associated with the perturbation optimization. Overall, the proposed scheme was proven to offer a favourable performance-complexity tradeoff compared to VP, while securing the required performance for the mobile users.

7. REFERENCES

[1] M. Costa, “Writing on dirty paper,”*IEEE Trans. Inf. Theory,* vol. IT-29, pp. 439441, May 1983.

[2] U. Erez, S. Shamai and R. Zamir, “Capacity and lattice strategies for cancelling known interference,” *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3820-3833, Nov. 2005.

[3] C. Windpassinger, R. Fischer, T. Vencel, J. Huber, “Precoding in multiantenna and multiuser communications,” *IEEE Trans. on Wireless Comms.*, vol. 3, no. 4, pp.1305-1316, July 2004.

[4] C. Masouros, M. Sellathurai, T. Rantarajah, “Interference Optimized Tomlinson-Harashima Precoding for Transmit Power Reduction”, *IEEE Trans. Sig. Proc, in press*

[5] C. B. Peel, B. M. Hochwald, A. L. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication part I: channel inversion and regularization,” *IEEE Trans. Coms.*, vol. 53, no. 1, pp. 195-202, Jan 2005.

[6] C. Masouros, E. Alsusa, “Dynamic Linear Precoding for the Exploitation of Known Interference in MIMO Broadcast Systems,” *IEEE Trans. on Wireless Comms.*, vol 8, issue 3, pp. 1396-1404, March 2009.

[7] C. Masouros, “Correlation Rotation Linear Precoding for MIMO Broadcast Communications”, *IEEE Trans. on Sig. Proc.*, vol 59, issue 1, pp. 252-262, Jan 2011

[8] C. B. Peel, B. M. Hochwald, A. L. Swindlehurst, “A vector-perturbation technique for near-capacity multiantenna multiuser communication part II: perturbation,” *IEEE Trans. Coms.*, vol. 53, no. 3, pp. 537-544, Mar 2005.

[9] B. Hassibi and H. Vikalo, “On the sphere-decoding algorithm I. Expected complexity,” *IEEE Trans. Sig. Proc.,* vol. 53, no. 8, pp. 28062818, Aug. 2005.

[10] E. Agrell, T. Eriksson, A. Vardy, K. Zeger, “Closest point search in lattices,” *IEEE Trans. Info. Theroy* , vol.48, no.8, pp. 2201-2214, Aug 2002

[11] H.S. Han, S.H. Park, S. Lee, I. Lee, “Modulo Loss Reduction for Vector Perturbation Systems,” *IEEE Trans. Coms.*, vol.58, no.12, pp.3392-3396, Dec 2010

[12] S.H. Park, H.S. Han, S. Lee, I. Lee, “A Decoupling Approach for Low-Complexity Vector Perturbation in Multiuser Downlink Systems,” *IEEE Trans. Wirel. Comms.,* vol.10, no.6, pp.1697-1701, June 2011

[13] U.P. Rico, E. Alsusa, C. Masouros, “A Fast Least-Squares Solution-Seeker Algorithm for Vector-Perturbation,” *Proc. IEEE GlobeCom 2008*, Nov 2008