Review and Performance Comparison of Finite-rate Feedback Strategies: Vector Quantization, Systematic Unitary Construction and Grassmannian Packing

Eunmo Kang

December 15, 2004

Joint work with Vasanthan Raghavan

Abstract

Channel knowledge at the transmitter can greatly improve the error probability of multi-input multi-output wireless communication system. Though receivers can obtain accurate channel estimates through training with pilot signals, transmitters can have only limited amount of the channel knowledge due to the limited rate of the feedback channel. One simple solution is the receiver quantizes the channel estimate and transfers it back to the transmitter. To reduce the feedback bits, better methods of beamforming codebook design based on Grassmannian packing and systematic unitary design were developed. In this report, we review these methods and compare the performance of each methods. Surprisingly the vector quantization outperformed others in simulation.

1 Introduction

Wireless communications employing multiple antennas at the transmitter and the receiver have attracted huge interest from academia and industry as a way to enable the users' ever-growing need for high data rate and reliable communication.

Telatar [1] and Foschini [2] have shown that capacity grows linearly with the smaller of the number of transmit antennas and receive antennas. Tarokh [3] showed that multiple antenna systems can be made resilient against channel fading through space-time coding. Space-time coding spreads information bits over multiple antennas (spatial dimension) and multiple channel uses (temporal dimension) so that multiple replicas of the same bits arrive at the receiver. Smartly designed space-time codes can achieve as many number of replicas or diversity order as the number of receive antennas multiplied by the smaller of the number of transmit antennas and the block length in rich scattering environment [4].

1.1 Space-Time Codes

Several important space-time coding techniques have been developed to exploit the potential of multiple antenna communication systems. Spatial multiplexing, transmitting independent data sub-streams over each antenna, achieves full rate, that is, one symbol per transmit antenna per channel use, but it sacrifices diversity order [5]. On the other hand, orthogonal space-time block codes (oSTBC) achieve full diversity sacrificing data rate by transmitting orthogonal streams over each antenna.

In order to fill the gap between these two schemes, linear dispersion code was developed [6]. Each symbol is multiplied by a basis matrix and then the sum of the scaled matrices are transmitted. The basis matrices are chosen by iteration to maximize the capacity, but it does not guarantee optimal error performance.

Recently developed threaded algebraic space-time codes aims at minimizing the error probability while achieving both full rate and full diversity [7]. By imposing a structure on the codeword matrices, it transforms the original optimization problem of finding multiple basis matrices for linear dispersion codes into a much simpler problem of finding an optimal scalar. Currently this code achieves the best error performance competing with perfect space-time block codes [8].

1.2 Feedback

Most space-time codes, including all of the codes above, assume i.i.d. quasi-static block fading channels. The quasi-static block fading model assumes that the channel is constant for the block length, or while a set of symbols are transmitted, then changes to an *independent* state during the transmission of the next symbols. In practice, next channel is temporally correlated with the previous channel due to the limited mobility of transmitters, receivers and scatterers.

Temporal correlation can be used to our advantage because it allows the receiver to feed back some information about the current channel to the transmitter so that the transmitter can adaptively change its coding and modulation scheme. About 20 dB power gain relative to nonadaptive transmission was achieved by varying rate and power for single antenna systems [9] and adaptive scheme in more realistic correlation scenario [1] was considered in [10]. For multiple antenna systems, [11] showed that transmitter aided by the feedback information can improve error performance through precoding, where optimum precoding matrix is obtained by iterative search. For two transmit antenna case, the SNR maximizing precoding matrix was derived in a closed form in [12].

1.3 Limited Feedback and Outline

However, every entry of channel matrix cannot be fed back due to the finite-rate of the feedback channel. Therefore, feedback schemes with small bits of feedback information have been researched by many researchers [13] [14] [15] [16] [17]. In this report, we review three feedback methods: vector quantization (section 3), systematic unitary construction (section 4) and grassmannian packing (section 5). In chapter 6, we compare the error performance of these methods through simulation.

2 System Model

We consider narrow-band complex baseband equivalent model of the multiple-input multiple-output (MIMO) channel:

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}$$

A transmitter equipped with M_t antennas and a receiver with M_r antennas are linked by wireless channel represented by the $M_r \times M_t$ channel matrix $\mathbf{H}$. Entry (r, t) of $\mathbf{H}$ denoted by $\mathbf{H}(r, t)$ represents the complex path gain between

[1] Exact temporal correlation function might be hard to obtain.

the t-th transmit and the r-th receive antenna. A set of information symbols is transmitted using the channel L times, coded in the $M_t \times L$ codeword matrix $\mathbf{X}$. The receiver observes a noise corrupted $M_r \times L$ matrix $\mathbf{Y}$.

We make the following assumptions. A channel matrix consists of i.i.d. Gaussian distributed entries, $\mathbf{H}(r,t) \sim \mathcal{CN}(0,1)$, and so does noise, $\mathbf{N}(r,l) \sim \mathcal{CN}(0,N_0)$. The codeword matrix is normalized to transmit with power P per channel use. With frequency division duplexing system in mind, channel is perfectly known at the receiver through, e.g. pilot-assisted training, but not at the transmitter. Feedback channel has no delay and is error-free for simplicity.

3 PEP Minimization Using Vector Quantization

Jongren in [18] tackled the problem of finding the precoding [2] matrix $\mathbf{W}$ that minimizes pairwise error probability (PEP) for orthogonal STBC when some feedback information is available at the transmitter . Specifically, an estimate $\hat{\mathbf{h}} = \mathbf{m}_{\mathbf{h}|\hat{\mathbf{h}}}$ of the true channel $\mathbf{h}$ and the mean squared error matrix $\mathbf{R}_{\mathbf{hh}|\hat{\mathbf{h}}}$ are assumed known at the transmitter. Then the following is known [18].

Theorem 1. *PEP of selecting* $\hat{\mathbf{X}}$ *when* $\mathbf{X}$ *is actually transmitted, denoted by PEP(*$\mathbf{X} \to \hat{\mathbf{X}}$*), is minimized by minimizing the following criterion:*

$$l(\mathbf{X},\hat{\mathbf{X}}) = \mathbf{m}^*_{\mathbf{h}|\hat{\mathbf{h}}}\mathbf{R}^{-1}_{\mathbf{hh}|\hat{\mathbf{h}}}\mathbf{\Psi}(\mathbf{X},\hat{\mathbf{X}})^{-1}\mathbf{R}^{-1}_{\mathbf{hh}|\hat{\mathbf{h}}}\mathbf{m}_{\mathbf{h}|\hat{\mathbf{h}}} - \log\det(\mathbf{\Psi}(\mathbf{X},\hat{\mathbf{X}})),$$

where $\mathbf{\Psi}(\mathbf{X},\hat{\mathbf{X}}) = (\mathbf{I}_{M_r} \otimes \mathbf{\Delta}\mathbf{\Delta}^H)/4N_0 + \mathbf{R}^{-1}_{\mathbf{hh}|\hat{\mathbf{h}}}$ *and* $\mathbf{\Delta} = \mathbf{X} - \hat{\mathbf{X}}$.

It is also shown that finding the optimal matrix $\mathbf{W}_o$ is a convex optimization problem which can be solved by interior point methods, for example. But in asymptotic cases, we have a simple closed-form solution of $\mathbf{W}_o$:

- If there is no channel knowledge at the transmitter or if SNR is high, $\mathbf{W}_o$ is an identity matrix or any unitary matrix.

[2] With precoding, the transmitter sends $\mathbf{WX}$ instead of $\mathbf{X}$.

- If the transmitter has perfect channel knowledge, beamforming is optimal, i.e.,

$$\mathbf{W}_o = \left[\begin{array}{cccc} \mathbf{v}_{max} & 0 & \cdots & 0 \end{array} \right], \tag{1}$$

where $\mathbf{v}_{max}$ is the right singular vector of $\mathbf{H} = \mathbf{U}\mathbf{D}\mathbf{V}^H$ corresponding to the biggest singular vector.

With above tools at hand, Jongren in [19] considered an optimal feedback strategy incorporating feedback delay, quantization errors and feedback channel errors. To distinguish the next channel and the current channel, we use $\mathbf{h}$ and $\tilde{\mathbf{h}}$, respectively.

The encoder or quantizer at the receiver observes current channel $\tilde{\mathbf{h}}$ and outputs an index $i = \text{enc}(\tilde{\mathbf{h}})$ of a codeword in the quantization codebook. After passing i through the imperfect discrete feedback channel denoted by $P_{j|i}$ the transmitter receives j and finally the decoder outputs $\hat{\mathbf{h}}(j) = \text{dec}(j)$. The encoder and decoder pair minimizing $\mathbb{E}[||\mathbf{h} - \hat{\mathbf{h}}(j)||^2]$ can be obtained by the Lloyd algorithm.

The last step is obtaining the mean $\mathbf{m}_{\mathbf{h}|j}$ and variance $\mathbf{R}_{\mathbf{hh}|j}$ of the true channel conditional on the received index j, in order to finally get $\mathbf{W}_o$. They are given by

$$\begin{aligned} \mathbf{m}_{\mathbf{h}|j} &= \hat{\mathbf{h}}(j) \\ \mathbf{R}_{\mathbf{hh}|j} &= \mathbf{R}_{\mathbf{hh}|\tilde{\mathbf{h}}} + \mathbf{R}_{\mathbf{h}\tilde{\mathbf{h}}}\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}^{-1}(\mathbb{E}[\tilde{\mathbf{h}}\tilde{\mathbf{h}}^H|j] - \mathbb{E}[\tilde{\mathbf{h}}|j]\mathbb{E}[\tilde{\mathbf{h}}|j]^H)\mathbf{R}_{\tilde{\mathbf{h}}\tilde{\mathbf{h}}}^{-1}\mathbf{R}_{\mathbf{h}\tilde{\mathbf{h}}}^H. \end{aligned} \tag{2}$$

Note that (1), (2) and the assumptions given in section 2 lead to the interesting fact – When feedback channel is ideal, oSTBC with precoding boils down to beamforming with the dominant right singular vector of the channel. Furthermore, if we have only one receive antenna which is the setup for the later simulation and comparison of the error performance, the beamforming vector is nothing but the channel matrix itself. Hence, the encoder-decoder pair for the feedback channel needs only to transmit the $1 \times M_t$ channel matrix with minimum squared error.

4 Outage Probability Minimization Using Systematic Unitary Construction

Instead of minimizing PEP, Mukkavilli *et al* looked at the finite-rate feedback problem from the perspective of outage probability. They considered a system with multiple transmit antennas and a single receive antenna, which transmits a complex data symbol s after being steered by a beamforming vector $\mathbf{c}_k$:

$$\begin{aligned} y &= \mathbf{h}^T \mathbf{x} + n \\ &= \mathbf{h}^T \mathbf{c}_k s + n. \end{aligned}$$

The outage probability of this system with target data rate R is given by

$$P_{out}(R, P) = \sum_{k=1}^{K} P(\mathbf{c}_k) P_{out}(R, P|\mathbf{c}_k),$$

where the outage probability conditioned on $\mathbf{h} \in H_k$ is

$$P_{out}(R, P|\mathbf{c}_k) = \Pr[| < \mathbf{h}, \mathbf{c}_k > |^2 < (2^R - 1)/P | \mathbf{h} \in H_i]$$

and H_k is the set of channel realizations which will be quantized to $\mathbf{c}_k$. Choosing $\mathbf{c}_k$ given a codebook is easy.

Lemma 2. *Given a beamformer codebook* $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_K\}$ *and a channel realization* $\mathbf{h}$*, the outage probability-minimizing* $\mathbf{c}_k$ *is the one that maximizes* $| < \mathbf{h}, \mathbf{c}_k > |$.

The construction of a codebook follows.

Theorem 3. *The design criterion for a beamformer is equavalent to the following optimization problem*

$$\begin{aligned} \max_{\mathcal{C} \in \mathbf{C}^{M_t}} \gamma^*(\mathcal{C}) &= \max_{\mathcal{C} \in \mathbf{C}^{M_t}} \frac{2\gamma_0}{1 + \max_{k \neq m} | < \mathbf{c}_k, \mathbf{c}_m > |} \\ &\leftrightarrow \min_{\mathcal{C} \in \mathbf{C}^{M_t}} \max_{k \neq m} | < \mathbf{c}_k, \mathbf{c}_m > | \end{aligned} \tag{3}$$

where γ_0 *is a minimum squared channel norm that achieves nonzero outage probability.*

There are two ways to solve this optimization problem: systematic unitary construction and Grassmannian packing. Here we look at the first method and the second one will be discussed in the section 5.

4.1 Algebraic Systematic Unitary Construction

Systematic unitary construction(SUC) method was developed in [20] to construct constellations for noncoherent space-time codes. Interestingly, design criterion for this problem is exactly same as the beamformer codebook design problem at hand if we set the block length L and M_t of the SUC problem to M_t and 1, respectively, of the beamformer problem.

Beamformer codebook $\{\mathbf{c}_k\}$ can be algebraically constructed as follows. Every arithmetic is modulo-q, where q is chosen to satisfy $q^N = K$ and usually $q = K, N = 1$. First, form $\mathbf{k} = \begin{bmatrix} \mathbf{k}_1 & \mathbf{k}_2 & \cdots & \mathbf{k}_N \end{bmatrix}$ where each $\mathbf{k}_n$ is in the ring of integers modulo-q, $\mathbb{R}_q$. Then $\mathbf{a}_k = \mathbf{kU}$ where $\mathbf{U}$ is a generator matrix whose entries are also in $\mathbb{R}_q$. Finally, $\mathbf{c}_k$ is formed by

$$\mathbf{c}_k = \frac{1}{\sqrt{M_t}} \begin{bmatrix} e^{i2\pi[\mathbf{a}_k]_1/q} \\ e^{i2\pi[\mathbf{a}_k]_2/q} \\ \vdots \\ e^{i2\pi[\mathbf{a}_k]_{M_t}/q} \end{bmatrix}.$$

Note that there is no deterministic algorithm to choose the optimal generator matrix $\mathbf{U}$, so we should rely on the iteration of random selection. However, this is not too much of burden because once the beamforming codebook is obtained off-line, there is no further need for realtime computation.

5 SNR Maximization Using Grassmannian Packing

Mukkavilli *et al.* [21] [15] and Love *et al.* [14] independently discovered that Grassmannian packing gives optimal solution for the finite-rate feedback problem, although they worked with different criterion, outage probability and SNR maximization, respectively.

In [14], authors considered the following system with beamforming and combining

$$\begin{aligned} y &= \mathbf{z}^H\mathbf{Y} \\ &= \mathbf{z}^H\mathbf{Hc}s + \mathbf{z}^H\mathbf{n} \end{aligned}$$

whose SNR γ is given by

$$\gamma = \frac{P|\mathbf{z}^H\mathbf{Hc}|^2}{||\mathbf{z}||^2 N_0}.$$

Maximum Ratio Combining(MRC) achieves the upperbound of

$$|\mathbf{z}^H\mathbf{Hc}|^2 \leq ||\mathbf{z}||^2||\mathbf{Hc}||^2$$

with $\mathbf{z} = \mathbf{Hc}/||\mathbf{Hc}||$ leading to

$$\gamma = \frac{P||\mathbf{Hc}||^2}{N_0},$$

so we always assume that the receiver uses MRC in this section.

If the transmitter has complete knowledge of the channel, the transmitter can use singular value decomposition to get $\mathbf{H} = \mathbf{UDV}^H$, and then by setting $\mathbf{c}$ to the dominant right singular vector $\mathbf{v}_{max}$, it can achieve $||\mathbf{Hc}|| = \lambda_{max}$. However, using quantized version of $\mathbf{H}$,

$$\mathbf{Q}_{\mathcal{C}}(\mathbf{H}) = \arg \max_{1 \leq k \leq K} ||\mathbf{Hc}_k||^2,$$

degrades SNR and we get the following distortion function

$$\begin{aligned} \mathbf{G}(\mathcal{C}) &= \mathbb{E}_{\mathbf{H}}\left[\lambda_{max} - \sum_{t=1}^{M_t} \lambda_t |\mathbf{v}_t^H \mathbf{Q}_{\mathcal{C}}(\mathbf{H})|^2\right] \\ &\leq \mathbb{E}_{\mathbf{H}}\left[\lambda_{max}\right] \mathbb{E}_{\mathbf{H}}\left[1 - |\mathbf{v}_{max}^H \mathbf{Q}_{\mathcal{C}}(\mathbf{H})|^2\right] \end{aligned}$$

It is shown in [14] that the upperbound of $\mathbf{G}(\mathcal{C})$ is minimized by the following optimization:

$$\max_{\mathcal{C}} \delta(\mathcal{C}) = \max_{\mathcal{C}} \min_{1 \leq k < m \leq K} \sqrt{1 - |\mathbf{c}_k^H \mathbf{c}_m|^2} \quad (4)$$

$$\leftrightarrow \min_{\mathcal{C}} \max_{1 \leq k < m \leq K} |\mathbf{c}_k^H \mathbf{c}_m|^2 \quad (5)$$

Note that two criterion for outage probability (3) and SNR maximization (5) are exactly same. Noting that the square root term in (4) is exactly the definition of the chordal distance, we can use the rich results on Grassmannian packing for codebook design.

6 Simulation

We simulated and compared symbol error probability of three feedback strategies reviewed in this report. The system consists of 4 transmit antennas and

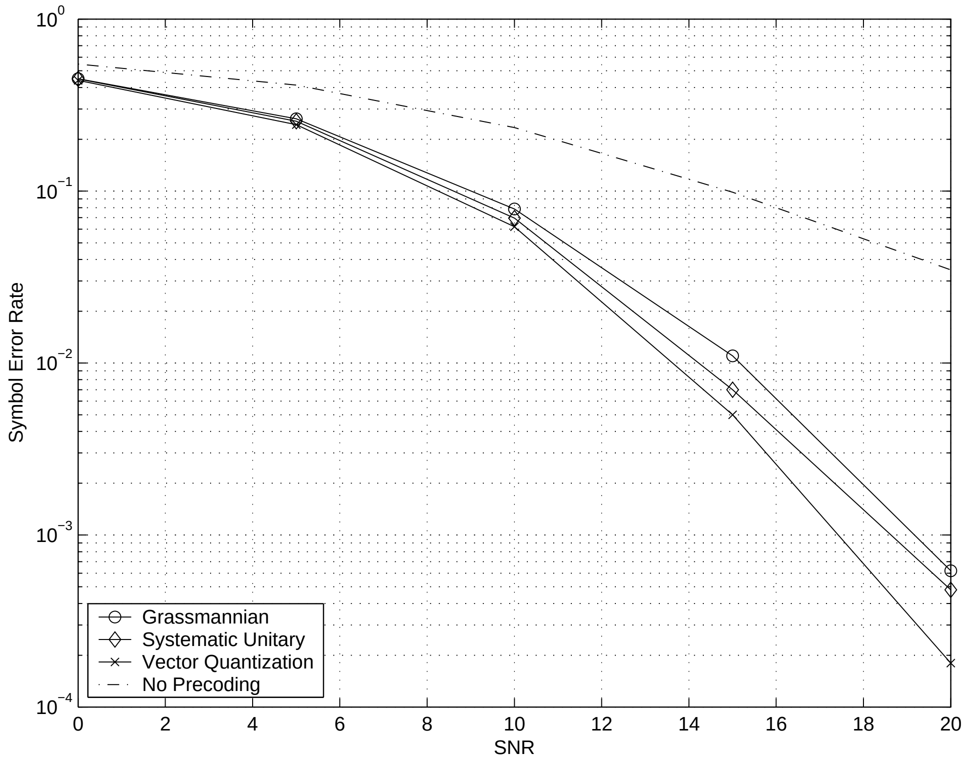

Figure 1: Symbol error rate comparison with M_t=4, M_r=1, 4-QAM, and cardinality of codebook 16

1 receive antenna. Based on the channel realization $\mathbf{h}$, beamforming matrix $\mathbf{c}_k$ is chosen from the codebook of size 16 according to Lemma 2, which is, choosing $\mathbf{c}_k$ such that $| < h, \mathbf{c}_k > |$ is maximized. This beamformer is multiplied by a data symbol s randomly drawn from 4-QAM constellation. The receiver uses maximum likelihood decoding.

The Grassmannian line packing matrices were downloaded from Prof. David Love's homepage, http://dynamo.ecn.purdue.edu/ djlove/, but for the other two schemes, we actually computed the beamformer codebooks using vector quantization and systematic unitary construction in a way discussed in this report. Many generator matrices $\mathbf{U}$ [3] for the systematic unitary construction produced same objective value given in (3).

Surprisingly, vector quantization which is the simplest and most straight-

[3] For example, [1 14 0 5] and [1 7 6 10]

forward method outperformed the other methods. Grassmannian packing which uses the established packing results were even beat by the systematic unitary construction.

As expected, feedback based communication achieved significant error performance improvement using only one bit per transmit antennas.

7 Conclusion

We have reviewed three methods of constructing beamformer codebooks to implement feedback with finite-rate feedback channels. Although each method targeted different criterion, e.g. PEP, outage performance, and SNR, all three methods boiled down to one criterion, placing each beamforming vector as far as possible. All three methods achieved significant improvement of 5-10 dB in error performance over no-feedback scheme.

There remains an interesting problem, why high profile Grassmannian packing was outperformed by simpler and iteration-based methods. Another direction of future study is the beamformer codebook design for correlated channels and multi-stream transmission.

References

[1] E. Telatar. Capacity of multi-antenna gaussian channels. *AT&T-Bell Labs Inter. Tech. Memo*, 1995.

[2] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Laboratories Technical Journal*, 1(2):41–59, 1996.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Trans. Inform. Theory*, 44(2):744–765, Mar. 1998.

[4] E. Kang, Z. Hong, and A.M. Sayeed. Space-time coding for correlated fading channel via virtual channel representation. *The 41st Allerton Conference on Communication, Control and Computing*, October 2003.

[5] P.W. Wolniansky, G.J. Foschini, G.D. Golden, and R.A. Valenzuela. V-blast: An architecture for realizing very high data rates over the rich-scattering wireless channel. *Proc. ISSSE-98*, September 1998.

[6] Babak Hassibi and Bertrand M. Hochwald. High-rate codes that are linear in space and time. *IEEE Trans. Inform. Theory*, 48:1804–1824, July 2002.

[7] H. El Gamal and M. O. Damen. Universal space-time coding. *IEEE Trans. Inform. Theory*, 49:1097–1119, May 2003.

[8] Frederique Oggier, Ghaya Rekaya, Jean-Claude Belfiore, and Emanuele Viterbo. Perfect space time block codes. *The 42st Allerton Conference on Communication, Control and Computing*, October 2004.

[9] Andrea J. Goldsmith and Soon-Ghee Chua. Variable-rate variable-power mqam for fading channels. *IEEE Trans. on Communications*, 45(10):1218–1230, October 1997.

[10] Dennis L. Goeckel. Adaptive coding for time-varying channels using outdated fading estimates. *IEEE Trans. on Communications*, 47(6):844–855, June 1999.

[11] G. Jongren, M. Skoglund, and B. Ottersten. Combining beamforming and orthogonal space-time block coding. *IEEE Trans. on Information Theory*, 48:611–627, March 2002.

[12] Shengli Zhou and G.B. Giannakis. Adaptive modulation for multi-antenna transmissions with channel mean feedback. *IEEE Trans. on Wireless Communications*, 3:1626–1636, September 2004.

[13] George Jongren and Mikael Skoglund. Quantized feedback information in orthogonal space-time block coding. *IEEE Trans. on Information Theory*, 50(10), October 2004.

[14] David J. Love and Robert W. Heath. Grassmannian beamforming for multiple-input multiple-output wireless systems. *IEEE Trans. on Information Theory*, 49(10), October 2003.

[15] K.K. Mukkavilli, A. Sabharwal, E. Erkip, and B. Aazhang. On beamforming with finite rate feedback in multiple-antenna systems. *IEEE Trans. on Information Theory*, 49(10), October 2003.

[16] D.J. Love and R.W. Heath. Limited feedback precoding for spatial multiplexing systems using linear receivers. *Military Communications Conference*, 1:627–632, October 2003.

[17] June Chul Rho and B.D. Rao. An efficient feedback method for mimo systems with slowly time-varying channels. *Wireless Communications and Networking Conference*, 2:21–25, March 2004.

[18] George Jongren, Mikael Skoglund, and Bjorn Ottersten. Combining beamforming and orthogonal space-time block coding. *IEEE Transactions on Information Theory*, March 2002.

[19] Jongren G. and Skoglund M. Utilizing quantized feedback information in orthogonal space-time block coding. *Proceedings IEEE Global Telecommunicaations Conference*, November 2000.

[20] Bertrand M. Hochwald, Thomas L. Marzetta, Thomas J. Richardson, Wim Sweldens, and Rudiger Urbanke. Systematic design of unitary space-time constellations. *IEEE Trans. on Information Theory*, 46(6):1962–1973, September 2000.

[21] K.K. Mukkavilli, A. Sabharwal, B. Aazhang, and E. Erkip. Performance limits on beamforming with finite rate feedback for multiple antenna systems. *The Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, 1:536–540, November 2002.