Feedback Reduction of Spatially Multiplexed MIMO Systems Using Compressive Sensing

Raymundo Nogueira de Sá Netto and Charles Casimiro Cavalcante

***Abstract*— In this paper we analyze spatially multiplexed MIMO systems with limited Channel State Information (CSI) and zero forcing (ZF) linear signal detection technique. Two schemes were considered: Quantization Codebook (QC) and Compressive Sensing (CS). Compressive Sensing is used to generate a reduced CSI feedback to the transmitter in order to reduce feedback load into the system. Performance of the schemes are compared by computational simulations of bit error rate (BER) curves for the considered approaches QC and CS.**

***Keywords*— BER, limited feedback, compressive sensing, quantization codebook.**

I. Introduction

Spatially multiplexed MIMO (SM-MIMO) systems can transmit data at higher speed than MIMO systems using antenna diversity techniques [1]. Zero Forcing (ZF) is one of the techniques we could use for signal detection in those systems. It needs complete Channel State Information (CSI) knowledge and a way to acquire it is using a feedback link from the receiver. However, even if the CSI can be perfectly estimated at the receiver, the required bandwidth for feedback is aggravated as the number of transmit and receive antennas increases [2]. Limited feedback beamforming is used to reduce the required bandwidth. When implemented, the beamforming vector is restricted to lie in a finite set or codebook that is known to both the transmitter and receiver [3].

Recently, Compressive Sensing (CS), also known as compressed sensing or compressive sampling, has been applied in diverse contexts of signal processing and communications, where the information content is sparse [3]. The spatial correlation between antenna arrays was exploited in [2] in order to obtain sparse representations of the channel and use CS to reduce the feedback load. CS is used in [4] to reduce the requirement of memory and complexity as the feedback rate increases and achieve greater sum throughput compared to Vector Quantization Codebook (VQC). In [5] it was proposed the use of CS to reduce the feedback for digital and analog schemes to achieve the same sum-rate throughput as the one achieved by dedicated feedback schemes, with limited feedback channels. Users that have a SINR larger than a threshold transmit the same feedback information and they are identified by the Base Station (BS) using CS and also a relation between sparsity and the threshold was given in [8]. In [9] a distributed self-selection procedure is combined with CS to identify a set of users who are getting simultaneous access to the downlink broadcast channel.

Raymundo Nogueira de Sá Netto and Charles Casimiro Cavalcante¸ Departamento de Engenharia de Teleinformática, Universidade Federal do Ceará, Fortaleza-CE, Brasil, E-mails: raymundo@gtel.ufc.br, charles@gtel.ufc.br.

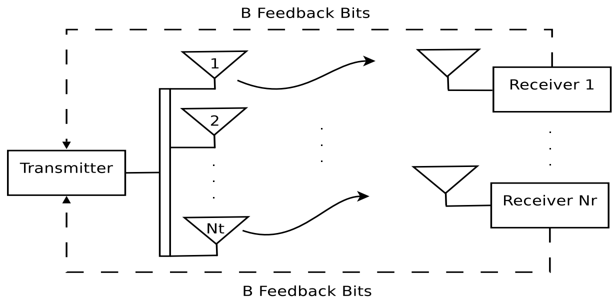

Fig. 1. Limited feedback system model

This work is based on [4], but instead of comparing the sum throughput between CS and VQC feedbacks, it uses CS to reduce the feedback load and to do a comparison between CS or VQC, both with limited feedback, and ZF technique with full CSI. This comparison has been made for two spatially multiplexed MIMO systems and the element of analysis is the Bit Error Rate (BER) over a Signal-to-Noise Ratio (SNR) variation.

The remainder of this paper is organized as follows. Section II provides the system model, as well as a review of CS operation. Section III shows how the feedback load reduction occurs. The results obtained by the two feedback protocols are shown in Section IV. Finally, in Section V, our conclusions are stated.

II. System Model

We consider MIMO wireless communication system with N_t transmit antennas and N_r receive antennas, as shown in Figure 1. The received signal vector at the N_r antennas is written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{H} = [\mathbf{h_1}, \mathbf{h_2}, ..., \mathbf{h}_{N_t}]^T$ is the channel matrix with independent and identically distributed (i.i.d) complex zero-mean unit variance Gaussian random values, $\mathbf{n}$ is an additive white Gaussian noise (AWGN) and $\mathbf{x}$ is the precoded vector that satisfies an average transmit power constraint $E\{\mathbf{x^H x}\} = 1$, where $E = \{.\}$ is the expectation operator and $()^H$ is the conjugate transpose..

One way to calculate the BER for the MIMO system considered at the reception by (1) is the usage of a ZF signal scheme [4], [5]. This is a linear signal detection method and treats all transmitted signals as interferences except for the desired stream from the target transmit antennas. To ease the detection of desired signals from each antenna, the effect of the channel is inverted by a weight matrix $\mathbf{W}$ such that

$$\mathbf{y} = \mathbf{HWx} + \mathbf{n}, \tag{2}$$

where $\mathbf{W}$ in ZF technique is defined by

$$\mathbf{W} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H, \tag{3}$$

where $(.)^H$ denotes the Hermitian transpose operation. $\mathbf{W}$ is calculated at the receivers and they feed it back to the transmit antennas, i.e., $\mathbf{W}$ represents the CSI. For our purposes it has been considered that $\mathbf{W}$ can be perfectly estimated at the receivers and the feedback channel is noiseless and delay free. However, as shown in Figure 1, the feedback channel is limited, so the CSI is not completely obtained by the transmit antennas. Thus, BER is not the same as the one considering ZF technique with full CSI. In this work, we assume the limited feedback channel has two protocols being used: quantization codebook (QC) and compressive sensing.

A. Quantization Codebook

In a limited feedback channel each receive antenna quantizes its channel using B bits and feeds back these bits. The quantization is performed using a vector quantization codebook that is known at the transmit and receive antennas. Typically, each receive antenna uses a differente codebook to prevent multiple antennas from quantizing their channel to the same quantization vector [4]. A quantization codebook $\mathbf{C} = \{\mathbf{c}_1, ..., \mathbf{c}_{2^B}\}$ consists of 2^B codeword vectors. Each codeword vector $\mathbf{c_i}$ has unit norm and length equal to the number of transmit antennas N_t.

The receive antenna quantizes its channel to the quantization vector (codeword) that is closest to its channel vector. This closeness is measured in terms of the angle between two vectors or, equivalently, the inner product [4]. Thus, a receive antenna i obtains the quantization index F_i according to [3]

$$\begin{aligned} F_i &= \arg \max_{j=1,...,2^B} |\mathbf{h}_i^H \mathbf{c}_j| \\ &= \arg \min_{\mathrm{j}=1,...,2^{\mathrm{B}}} \sin^2\left(\angle(\mathbf{h}_{\mathrm{i}}, \mathbf{c}_{\mathrm{j}})\right), \end{aligned} \tag{4}$$

and feeds this index back to the transmit antenna. The choice of vector quantization codebook significantly affects the quality of the CSI provided to the transmit antenna, i.e., the larger is the codebook the better is the quality of the CSI.

B. Compressive Sensing

Before explaining how CS is used on a limited feedback scheme, a brief overview is shown. This emerging theory is based on exploiting the sparsity present into the signals, being able to recover these signals from a limited number of linear measurements, and it is more effective compared to the classical Nyquist-Shannon sampling [5], [11]. Sparsity expresses the idea that the "information rate" of a continuous time signal may be much smaller than suggested by its bandwidth, or that a discrete-time signal depends on a number of degrees of freedom which is comparably much smaller than its (finite) length. More precisely, CS exploits the fact that many natural signals are sparse or compressible in the sense that they have concise representations when expressed in a proper basis $\mathbf{A}$ [6].

Let $\mathbf{x}$ be a $N \times 1$ vector, with at most S non-zero elements, where $S << N$. Consider a $M \times N$ measurement matrix $\mathbf{A}$, where $M << N$ and $M > S$. The measurements can be obtained by

$$\mathbf{b} = \mathbf{Ax}, \tag{5}$$

where $\mathbf{b}$ is a $M \times 1$ vector. Since $M << N$ the vector $\mathbf{x}$ can be represented by $\mathbf{b}$ with much less information, i.e., $\mathbf{x}$ is compressed into $\mathbf{b}$.

Hence, this system has more unknowns than equations, and thus it has either no solution, if $\mathbf{b}$ is not in the span of the columns of the matrix $\mathbf{A}$, or infinitely many solutions. To avoid these conditions to happen, $\mathbf{A}$ has to have l_2-normalized columns, and for an integer scalar $s \leq n$, consider submatrices A_s containing s columns from $\mathbf{A}$. δ_s is defined as the smallest quantity such that

$$\forall \mathbf{x} \in \mathbf{R}^s : (1-\delta_s)||\mathbf{x}||_2^2 \leq ||\mathbf{A}_s\mathbf{x}||_2^2 \leq (1+\delta_s)||\mathbf{x}||_2^2, \tag{6}$$

holds true for any choise of s columns. Then $\mathbf{A}$ is said to have an *s-restricted isometry property* (RIP) with a constant δ_s [11]. This constant measures how orthonormally close the column vectors of the measurement matrix $\mathbf{A}$ are to each other.

On the other side, a recovery algorithm has to be used to obtain $\mathbf{x}$ from $\mathbf{b}$. If the RIP holds, then the following linear program gives an accurate reconstruction [10]:

$$\min_{\mathbf{x} \in \mathbf{R}^n} ||\mathbf{x}||_{l_1} \text{ subject to } \mathbf{Ax} = \mathbf{b}. \tag{7}$$

To solve this optimization task, there are some proposed algorithms on literature: *l_1-Magic* [12], *Orthogonal Matching Pursuit* (OMP) [13], *Basis Pursuit* (BP) [14], *Dantzig-Selector* (DS) [11].

Returning to our purpose, we desire to compress the CSI to reduce the feedback load. According to $\mathbf{H}$ features, it is not sparse to use CS for compression. Song *et al.* proposed a CSI sparse approximation method [4]. Like the quantization codebook, a $N_t \times 2^B$ matrix $\mathbf{Q} = [\mathbf{q_1} \dots \mathbf{q_m} \dots \mathbf{q_{2^B}}]$, called dictionary, is required. After this, the method consists of three steps:

1) Column Selection: S columns maximally correlated with their own CSI ($\mathbf{h}$) are selected. Let $\pi(i)$ be the ith column index largely correlated with the channel vector. Each receive antenna selects the column maximally correlated from the initial column index set $\Lambda_0 = \{1, ..., m, ..., 2^B\}$ of $\mathbf{Q}$ as $\pi(1) = \arg\max_{m \in \Lambda_0} |\langle \mathbf{q}_m, \mathbf{h} \rangle|$, where $\langle , \rangle$ stands for the inner product, and $|.|$ stands for the absolute value. After selecting $i-1$ columns, the ith column can be selected within $\Lambda_{i-1} = \{m | m \in \Lambda_{i-2} \text{ and } m \neq \pi(i-1)\}$ as $\pi(i) = \arg\max_{m \in \Lambda_{i-1}} |\langle \mathbf{q}_m, \mathbf{h} \rangle|$. This continues until the S columns are selected. $\Pi(i) = \{\pi(1), ..., \pi(S)\}$ is the index set for the selected columns. Since this procedure can find the S-dimensional subspace maximally correlated to $\mathbf{h}$ over the N_t-dimensional complex domain, the approximation error can be minimized.

2) CSI Approximation: Each receive antenna channel vector $\mathbf{h}$ is aproximated by the selected S columns in the minimum-mean-square-error sense, as follows:

$$\begin{aligned}\mathbf{z}(\Pi) &= \arg\min_{\mathbf{a}(\Pi)} ||\mathbf{h} - \mathbf{Q}(\Pi)\mathbf{a}(\Pi)||^2 \\ &= \mathbf{Q}(\Pi)^{\dagger}\mathbf{h} = (\mathbf{Q}(\Pi)^{\mathrm{H}}\mathbf{Q}(\Pi))^{-1}\mathbf{Q}(\Pi)^{\mathrm{H}}\mathbf{h},\end{aligned} \quad (8)$$

where $||.||$ is the *Frobenius* norm, $()^{\dagger}$ is the pseudoinverse, $\mathbf{Q}(\Pi)$ is the $N_t \times S$ matrix consisting of column vectors corresponding to Π, $\mathbf{z}(\Pi)$ is the optimal coefficient vector for $\mathbf{Q}(\Pi)$, and $\mathbf{a}(\Pi)$ is a candidate vector for $\mathbf{z}(\Pi)$.

3) Sparse CSI creation: The $2^B \times 1$ sparse CSI can be obtained as follows:

$$\tilde{\mathbf{h}} = \mathbf{Q}(\Pi)\mathbf{z}(\Pi) = \mathbf{Q}\mathbf{z}. \quad (9)$$

In the first equality, the fact that $\mathbf{Q}(\Pi)$ and $\mathbf{z}(\Pi)$ depend on Π makes it difficult to design a universal compression matrix, which does not depend on the channel vector. To design the given universal compression matrix, we must redefine $\tilde{\mathbf{h}}$ in the second equality of (9) with $\mathbf{Q}$ being independent of Π and $\mathbf{z}$ only containing S nonzero elements corresponding to Π.

Thus, $\mathbf{z}$ is S-sparse and CS can be used to compress this information to be feedback to the transmit antennas

$$\mathbf{b} = \mathbf{A}\mathbf{z}, \quad (10)$$

where $\mathbf{A}$ is a $M \times 2^B$ measurement matrix, $\mathbf{z}$ can be recovered by one of the algorithms previously mentioned, at the transmit antennas, and $\mathbf{h}$ can be estimated by $\mathbf{h} = \mathbf{Q}\mathbf{z}$.

So, the difference between quantization codebook and compressive sensing is that the transmit and receive antennas in quantization codebook has to know the codebook $\mathbf{C}$ used at each antenna. On the other hand, in compressive sensing they has to know the dictionary $\mathbf{Q}$ and the measurement matrix $\mathbf{A}$.

III. Feedback Load Reduction

For our studies, it was considered two spatially multiplexed MIMO systems. The first one has $N_t = N_r = 4$ and the second $N_t = N_r = 8$. The sparsity value $S = 1$ was the same for both systems. When we simulated the $N_t = N_r = 8$ system, to sparsify the 8×8 channel matrix $\mathbf{H}$ we had to divide it into four 4×4 matrices to use the same sparsity value and the number of measurements M as the first system ($N_t = N_r = 4$) and obtain the same feedback load reduction. S can be recovered with high probability as long as M is sufficiently large and $\mathbf{A}$ satisfies RIP [10]. Since $\mathbf{A}$ is an i.i.d Gaussian matrix, according to [8] M has to satisfy

$$M \geq kS\log_2(N_t/S), \quad (11)$$

where k is a constant. Considering $k = 1$, we obtain $M = 2$, which is used at all simulations.

Defined these parameters, the feedback load reduction is calculated for the two systems utilized. We have the following example that works for every value of B bits:

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{12^B} \\ a_{21} & a_{22} & \ldots & a_{22^B} \end{bmatrix} \begin{bmatrix} 0_{11} \\ \vdots \\ z \\ \vdots \\ 0_{12^B} \end{bmatrix}^T = \begin{bmatrix} b_{11} \\ b_{21} \end{bmatrix}, \quad (12)$$

where (12) is the matrix representation of (10). Since $S = 1$, $\mathbf{z}$ contains only one non-zero element and this sparse representation can be compressed into $M = 2$ measurements, equal to the length of $\mathbf{b}$. Considering $N_t = N_r = 4$, it means that $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \mathbf{h}_4]^T$ and each $\mathbf{h}$ can be compressed into $\mathbf{b}$. So the necessary feedback load for this case is represented by the following matrix representation:

$$\mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{bmatrix}. \quad (13)$$

This feedback load is smaller than the one used by the quantization feedback, since each vector $\mathbf{h}$ has to be representated by one codeword $\mathbf{c}$ and its length is N_t. So, the feedback load necessary by quantization feedback, in this case, is 32 and the one necessary by compressive sensing is 16, which is represented by each value of (13). It means that, utilizing compressive sensing, the feedback load can be reduced to half of the one used by quantization feedback. It is according to the compression ratio equation provided by [2]:

$$\eta = T_M/(N_T N_R), \quad (14)$$

where T_M represents the total number of measurements. The same compression ratio (η) is obtained when $N_t = N_r = 8$, since, as explained before, each receive antenna separate in two 4×1 vectors the 8×1 CSI vector calculated at it.

IV. Results

The results were obtained considering the ZF technique model using 4-QAM modulation provided by [1] with some adaptation to the feedback limitation and it also provided the codebook design parameters which are the same in IEEE 802.16e specification. All results show the BER variation as SNR increases between the two protocols discussed before for limited feedback and the same variation for ZF, but with full CSI. Considering $B = 10$ bits, $N_t = N_r = 4$ and the other parameters already defined at section III , the first result is shown in Figure 2.

In order to evaluate if the feedback load reduction was going to harm the BER, Figure 2 shows that, for a dictionary in CS with the same length of the codebook in QC, there is no degradation in terms of BER, i.e., they show the same behavior. Comparing both limited feedback schemes to ZF with full CSI, their behaviors is very different. At ZF as long as SNR increases, BER decreases. Otherwise, at the two schemes, the decaiment of BER is worst than the first and it tends to stabilize even if SNR increases.

In order to achieve a better performance for the two limited feedback schemes compared to ZF with full CSI and still compare all of them, there was a raise in the dictionary and

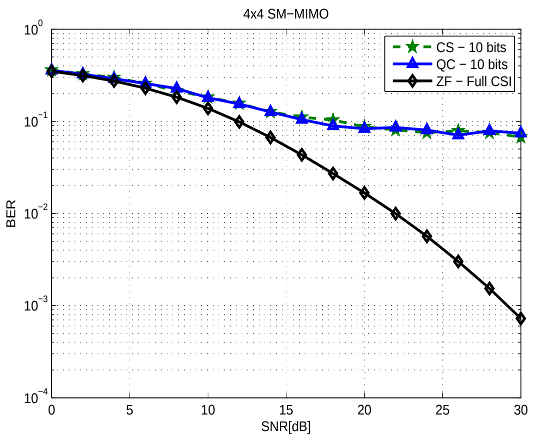

Fig. 2. BER vs SNR for the three schemes utilized with $B = 10$ bits and 4×4 SM-MIMO.

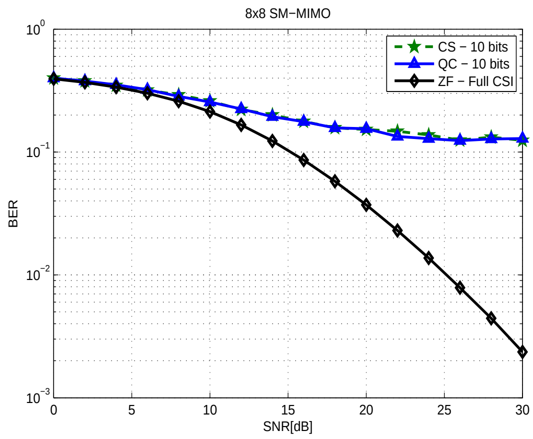

Fig. 4. BER vs SNR for the three schemes utilized with $B = 10$ bits and 8×8 SM-MIMO.

codebook lengths. This second result was obtained utilizing $B = 13$ bits and is shown in Figure 3.

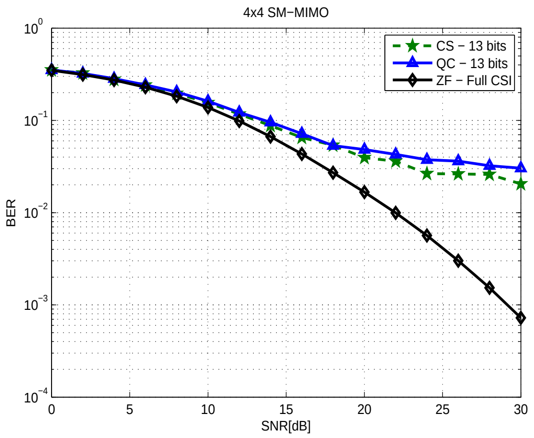

Fig. 3. BER vs SNR for the three schemes utilized with $B = 13$ bits and 4×4 SM-MIMO.

Figure 3 shows that utilizing $B = 13$ bits a better approximation to ZF with full CSI is achieved, but, even with this performance increase, the limited feedback schemes continue tending to stabilize, at this time with lower BER and higher SNR, showing the advantage of using more bits. This tendency continues for every raise in the number of bits B. Comparing only CS and VQ, they obtain the same BER variation.

From now on, the results obtained were considering $N_t = N_r = 8$, and the other parameters were kept the same. So, for $B = 10$ and $B = 13$ bits the relation BER vs SNR is shown, respectively, in Figures 4 and 5.

In Figures 4 and 5 BER exhibit the same behavior presented on Figures 2 and 3, respectively. Comparing CS and QC, BER variation is the same and increasing the number of bits, it was obtained a better approximation to ZF with full CSI for both limited feedback schemes.

The advantage of using CS is shown over all the results over the obtained feedback load reduction in terms of BER compared to QC, since they always have the same BER

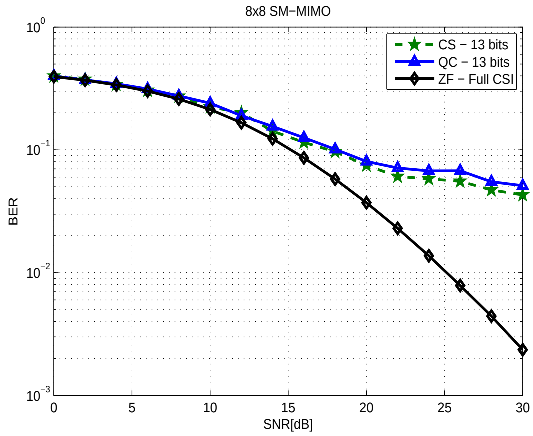

Fig. 5. BER vs SNR for the three schemes utilized with $B = 13$ bits and 8×8 SM-MIMO.

behavior. Otherwise, the amount of required memory and the computational complexity, as shown in Table I, is higher than in QC. Note that one complex value is saved by a double-precision floating point (64bits/value).

V. CONCLUSIONS

This paper suggests using CS to reduce the load of CSI feedback in an i.i.d SM-MIMO systems. The feedback compression can be carried out by compressive sensing, and highly-accurate channel information recovery can be achieved with a significant compression ratio, in our case 50%. Obtained the feedback load reduction, the performance was not compromised compared to quantization codebook scheme. So, we conclude that CS is a promising approach to reduce CSI feedback for SM-MIMO systems. In order to implement the utilized mechanisms in real systems, many practical issues that we did not consider, such as channel estimation errors and radio impairments in the feedback link, should be adressed in future studies.

TABELA I

COMPARISON OF THE AMOUNT OF REQUIRED MEMORY AND THE COMPUTATIONAL COMPLEXITY.

	Compressive Sensing	Quantization Codebook
Amount of required memory	$\underbrace{64 \times N_t \times 2^B}_{\text{dictionary}} + \underbrace{64 \times \mathrm{C} \times 2^{\mathrm{B}}}_{\text{measurement matrix}}$	$\underbrace{64 \times N_t \times 2^B}_{\text{codebook}}$
Computational complexity	$\underbrace{S \times N_t \times 2^B}_{\text{max. correlation}} + \underbrace{S \times C}_{\text{compression}} + \underbrace{N_t^3 + S(N_t^2 + N_t)}_{\text{sparse approximation}}$	$\underbrace{N_t \times 2^B}_{\text{max. correlation}}$

ACKOWNLEDGMENT

This work was supported by the Brazilian agencies CNPq and FUNCAP/PRONEX.

REFERÊNCIAS

[1] Y. S. Cho, J. Kim, W. Y. Yang and C. Kang, *MIMO-OFDM Wireless Communications with MATLAB*. Wiley, 2010.

[2] P. Kuo, H. T. Kung and P. Ting, *Compressive Sensing Based Channel Feedback Protocols for Spatially-Correlated Massive Antenna Arrays*. IEEE Wireless Communications and Networking Conference (WCNC), 2012.

[3] C. K. Au-Yeung and D. J. Love, *On the Performance of Random Vector Quantization Limited Feedback Beamforming in a MISO System*. IEEE Transactions on Wireless Communications, 2007.

[4] H. Song, W. Seo, and D. Hong, *Compressive Feedback Based on Sparse Approximation for Multiuser MIMO Systems*. IEEE Transactions on Vehicular Technology, 2007.

[5] S. T. Qaseem and T. Y. Al-Naffouri, *Compressive Sensing for Feedback Reduction in MIMO Broadcast Channels*. IEEE Transactions on Wireless Communications, April, 2009.

[6] E. J. Candes and M. B. Wakin, *An Introduction to Compressive Sampling*. IEEE Signal Processing Magazine, March, 2008.

[7] N. Jindal, *MIMO Broadcast Channels With Finite-Rate Feedback*. IEEE Transactions on Information Theory, November, 2006.

[8] K. Kim, S. Jang and D. Kim, *An Efficient Feedback Scheme Using Compressive Sensing for MIMO Broadcast Channel with Random Beamforming*. The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010

[9] S.R. Bhaskaran, L. Davis, A. Grant, S. Hanly and P. Tune *An Efficient Feedback Scheme Using Compressive Sensing for MIMO Broadcast Channel with Random Beamforming*. IEEE Information Theory Workshop on Networking and Information Theory, 2009. ITW 2009.

[10] E. Candes, J. Romberg and T. Tao, *Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information*. IEEE Trans. Inf. Theory, vol. 52, no. 2, pp. 489-509, Feb. 2006.

[11] M. Elad, *Sparse and Redundant Representations - From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[12] E. J. Candes and J. Romberg, *l1-MAGIC: Recovery of Sparse Signals via Convex Programming*, 2005. Available at http://www.acm.caltech.edu/l1magic.

[13] J. A. Tropp and A. C. Gilbert, *Signal recovery from partial information via Orthogonal Matching Pursuit*. IEEE Trans. Inform. Theory, April, 2005.

[14] S. S. Chen, D. L. Donoho, and M. A. Saunders, *Atomic decomposition by Basis Pursuit*. SIAM Review, 43(1):129-159, 2001.

[15] D. L. Donoho, *Compressed Sensing*. IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289-1306 April, 2006.

[16] J. Haupt, W. U. Bajwa, M. Rabbat and R. Nowak, *Compressed Sensing for Networked Data*. IEEE Signal Processing Magazine, March, 2008.