

On the Feedback Reduction of Relay Aided Multiuser Networks using Compressive Sensing

Khalil M. Elkhail, Mohammed E. Eltayeb, *Student Member, IEEE*, Abla Kammoun, *Member, IEEE*, Tareq Y. Al-Naffouri, *Member, IEEE*, and Hamid Reza Bahrami, *Member, IEEE*

Abstract—In this paper, we propose a feedback reduction scheme for full-duplex relay-aided multiuser networks. The proposed scheme permits the base station (BS) to obtain channel state information (CSI) from a subset of strong users under substantially reduced feedback overhead. More specifically, we cast the problem of user identification and CSI estimation as a block sparse signal recovery problem in compressive sensing (CS). Using existing CS block recovery algorithms, we first obtain the identity of the strong users and then estimate their CSI using the best linear unbiased estimator (BLUE). To minimize the effect of noise on the estimated CSI, we introduce a back-off strategy that optimally backs-off on the noisy estimated CSI and derive the error covariance matrix of the post-detection noise. In addition to this, we provide exact closed form expressions for the average maximum equivalent SNR at the destination user. Numerical results show that the proposed algorithm drastically reduces the feedback air-time and achieves a rate close to that obtained by scheduling schemes that require dedicated error-free feedback from all the network users.

Index Terms—Feedback, Decode-and-Forward, Full-Duplex Relaying, Compressive Sensing.

I. INTRODUCTION

In a relay aided networks, a relay station (usually low powered) is deployed to assist communication between the base station (BS) and a destination user. Relaying techniques can be classified, based on their forwarding strategy and required processing at the relay terminal, as *decode and forward* (DF) or *amplify and forward* (AF)

Khalil Elkhail, Abla Kammoun and Tareq Y. Al-Naffouri are with the Electrical Engineering Department, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia; e-mails: {khalil.elkhail, abla.kammoun, tareq.alnaffouri}@kaust.edu.sa. Tareq Y. Al-Naffouri is also associated with the Department of Electrical Engineering, King Fahd University of Petroleum and Minerals, Dhahran 31261, Kingdom of Saudi Arabia.

Mohammed Eltayeb is with the department of Electrical and Computer Engineering, The University of Texas at Austin, Texas, USA ; e-mail: meltayeb@utexas.edu.

Hamid Reza Bahrami is with the Department of Electrical and Computer Engineering, The University of Akron, Ohio, USA; e-mail: hrb@uakron.edu.

[5]. Relay transmission for both AF and DF relaying can be performed in *half-duplex* (HD) (e.g. [6]) or *full-duplex* (FD) (e.g. [7]) mode. In HD mode, the BS and the relay transmit on orthogonal channels, whereas in FD mode, the BS and relay share a common channel and the relay transmits and receives simultaneously over the same channel. Thus, half-duplex relay schemes induce a throughput loss due to the pre-log factor 1/2 [8]. Full duplex relay schemes are shown to improve the network throughput by eliminating this pre-log factor [7].

In a network with large number of users, the BS can exploit the multiuser diversity to maximize the network capacity. This is achieved by requesting all users to feed back their equivalent SNR (BS-relay-user) to the base station (BS). Clearly, this is expensive in terms of spectrum utilization and results in a large feedback overhead, especially when the number of users is high [9],[10]. However, channel state information (CSI) mismatch that could arise from both channel estimation errors and quantized feedback results in a loss in the achievable rate. Therefore, for practical scenarios, feedback strategies should be able to reduce the overhead while satisfying certain performance guarantees. In fact, to reduce the feedback, it is possible to request only a few strong users to encode their SNR on non-orthogonal codewords. This creates a sparse user regime, where sparse approximation algorithms (see e.g. [11]) are then applied to recover the identity and SNR of the strong users. The work in [16] uses Compressed Sensing (CS) to reduce the feedback load of single-input-single-output (SISO) networks, while other work e.g. [17]-[19] apply CS to reduce the feedback overhead for MIMO networks. In this paper, we apply CS to reduce the feedback overhead of a relay-aided multiuser network. We exploit the feedback nature of FD relays to pose the feedback problem as a block CS recovery problem. Block CS recovery leverages the a priori information of the signal block size to better differentiate true signal information from recovery artifacts. This leads to a more robust recovery [22]. To the best of our knowledge, this is the first paper that proposes a CS-based feedback algorithm for multi-user FD relay-aided networks. We focus on

networks that are comprised of a BS serving users in their respective coverage areas. It is assumed that there is no direct link between the BS and the users, and hence, communication must take place via a relay terminal. Examples of such a topology include cell-edge users, users that are shadowed from the BS, e.g. users in a shopping complex or an air-port, or users covered by a pico or a femto cell as the case in heterogeneous networks [20][21]. To exploit the multi-user diversity, the BS requires feedback of CSI from all the users. Based on the feedback the BS receives (via the relay), it schedules the user with the best BS-relay-user channel condition. This feedback requirement poses two immediate challenges. Firstly, CSI feedback (users-relay-BS) generates a great deal of feedback overhead (air-time) that could result in significant performance hits if the number of users becomes too high. Secondly, the feedback channel information is usually corrupted by additive noise. In this paper we address the above challenges and provide the following contributions

- We propose a CS-based full-duplex feedback algorithm that reduces the feedback overhead of multi-user relay aided networks. To the best of our knowledge, this is the first paper that introduces full-duplex feedback and accounts for the relay self-interference on the feedback link.
- We account for the feedback noise and refine the estimated equivalent SNR, and optimally *backing off*¹ on the noisy equivalent SNR estimates.
- We derive the post-CS detection noise variance.
- We derive the average equivalent SNR of the best user when the relay employs HD and FD feedback.

The remainder of the paper is organized as follows. In Section II, we introduce the system model. We describe the proposed feedback algorithm in Section III and evaluate its performance in Section IV. In Section V, we present some numerical results prior to concluding our work in Section VI.

Notations: Throughout this paper, we use the following notations : $F_X(\cdot)$, $F_X^{-1}(\cdot)$, $p_X(\cdot)$ and $\mathbb{E}(X)$ stand for the cumulative density function (CDF), the inverse CDF, the probability density function (PDF), and the expectation of the random variable X respectively. We denote by $\mathbb{P}(\mathcal{A})$, the probability of the event \mathcal{A} . Matrices are denoted by bold capital letters, rows and columns of the matrices are referred with lower case bold letters. We use the subscript c and r to refer to a matrix column and row respectively, for example $\mathbf{a}_{c,i}$ and $\mathbf{a}_{r,i}$

¹The term back-off denotes an SNR reduction, applied by the BS to the estimated SNRs, to reduce the likelihood of an SNR over-estimation error.

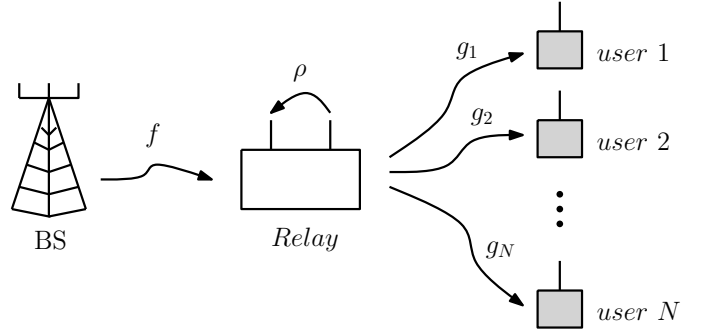


Figure 1. Network Model.

are respectively the i th column and row of the matrix \mathbf{A} . We denote by $\mathbf{A}_{\mathcal{I}}$, the submatrix of \mathbf{A} whose columns are in \mathcal{I} . $tr(\cdot)$ and $[\cdot]^t$ respectively denote the trace and the transpose operators on matrices. To refer to a matrix \mathbf{A} as a positive semi-definite matrix, we use the notation $\mathbf{A} \succeq 0$, consequently $\mathbf{A} \succeq \mathbf{B}$ if and only if $\mathbf{A} - \mathbf{B} \succeq 0$. We also denote by $\Gamma(\cdot)$, $\gamma(\cdot, \cdot)$ and $\Gamma(\cdot, \cdot)$, the Gamma function, the lower incomplete Gamma function and the upper incomplete Gamma function respectively.

II. SYSTEM MODEL

In this section, we introduce the downlink and feedback (uplink) transmission models for the relay-aided multi-user network.

A. Downlink Transmission Model

We consider a network with N cell-edge users served by one BS via a single FD relay (no direct path exists between the BS and destination users). As shown in Fig. 1, both the BS and the users are equipped with a single antenna, whereas the relay is equipped with one transmission antenna and one receive antenna. The two relay antennas operate simultaneously over the same frequency, i.e. reception and transmission occur at the same time over the same frequency. We assume that a direct link exists between the BS and the relay, hence, the channel from the BS to the relay f is modeled by a Rician distribution with fading parameter $K = \frac{b^2}{2\sigma^2}$, where $2\sigma^2$ is the average power in the Non-Line Of Sight (NLOS) multipath components and b^2 is the power in the LOS component [4]. For simplicity, the BS-relay channel can be approximated to follow a Nakagami fading channel parameterized by $P = b^2 + 2\sigma^2$ and fading parameter $d = \frac{(K+1)^2}{2K+1}$. Therefore,

$$F_{|f|^2}(x) = \frac{\gamma(d, \frac{x}{\theta})}{\Gamma(d)}, \quad (1)$$

where $\theta = \frac{P}{d}$. The channels from the relay to the N users, $(g_n)_{n=1,2,\dots,N}$ are assumed to be independent and

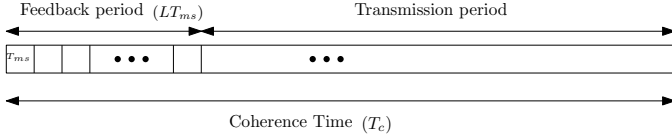


Figure 2. Structure of a frame of duration T_c composed of feedback period and transmission period.

identically distributed (i.i.d.) zero-mean complex Gaussian random variables with variance σ_g^2 . To maximize the downlink link rate, the destination user is selected based on the following rule

$$n^* = \arg \max_n \gamma_n, \quad (2)$$

where $\gamma_n = \frac{P_r |g_n|^2}{N_0}$ is the instantaneous receive SNR at the n th user which is an exponential random variable with mean $\bar{\gamma} = \frac{P_r \sigma_g^2}{N_0}$, and N_0 is the additive noise variance. The relay may apply cancellation techniques [24],[25] to (partly) eliminate this self-interference. Usually this involves estimation of the loop channel h and then subtraction of the interfering signal. However, in practice, complete interference cancellation is not achieved due to non-ideal channel estimation and signal processing. In this paper, we model the effective gain of the loop channel due to the residual interference after cancellation as a multiplicative factor ρ that remains constant during the transmission interval.² Let $s_{n^*}(t)$ and $\hat{s}_{n^*}(t)$ represent the original and the decoded symbol by the relay respectively, then after selecting the destination user n^* , the BS broadcasts the symbol $s_{n^*}(t)$ which is received by the relay as

$$y_r(t) = \sqrt{P_s} f s_{n^*}(t) + \rho \sqrt{P_r} \hat{s}_{n^*}(t-1) + z_r(t), \quad (3)$$

where P_s and P_r represent the BS and the relay transmission powers respectively, and the term z_r represent the zero mean variance N_0 additive Gaussian noise at the relay. During the second hop, the relay will decode and forward³ the symbol, and the n^* th user receives

$$y_{n^*}(t) = \sqrt{P_r} g_{n^*} \hat{s}_{n^*}(t) + z_{n^*}(t), \quad (4)$$

where $z_{n^*}(t)$ is the zero mean variance N_0 additive Gaussian noise at the n^* th user. The transmitted symbols are assumed to have the same constant normalized transmitted power over time, i.e. $\mathbb{E}[|s(t)|^2] = 1$.

²This assumption can be justified by the fact that relay's antennas are fixed during transmission.

³Throughout the paper, we only consider decode and forward as the downlink protocol.

B. Feedback Model

We assume a reciprocal TDD setup so that the BS-relay and the relay-user channels on the uplink are the same as those of the downlink. As shown in Figure 2, the feedback channel is assumed to be slotted, shared, and opportunistically accessed by the users. There are L feedback mini-slots and each feedback transmission is received over a duration of one feedback mini-slot time T_{ms} (see Figure 2). All channel coefficients, i.e. f , h and $(g_n)_{n=1,2,\dots,N}$, are assumed to be constant during the feedback and data transmission periods, (i.e. all channels are fixed during one coherence interval of duration T_c). Each user is assigned an ID which is a unique signature sequence vector of dimension $M \ll N$ for use on the slotted feedback channel. The choice of M is discussed in Section III-C. The signature sequence vectors, are drawn from the columns of a real Gaussian matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ (with zero mean and variance $\frac{1}{M}$ i.i.d. entries), and are deterministically assigned by the BS to the users.

Prior to feedback, the relay broadcasts a pilot from which the BS and the users estimate their channels with the relay. Only users having an instantaneous receive SNR (Please refer to eq. (2)) higher than a predetermined threshold γ_{th} encode (or multiply) their SNRs (relay-user SNR) with their signature sequence vectors and transmit the combination to the BS (via the relay) after applying proper uplink channel compensation, i.e. the n th user scales its transmission by $\frac{1}{g_n}$. The remaining users remain silent or effectively feed back a "0". Let $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_N]^t$ denotes the feedback sparse vector, where x_n is the feedback value of the n th user and it is given by

$$x_n = \begin{cases} \gamma_n & , \gamma_n > \gamma_{th} \\ 0 & , \text{otherwise.} \end{cases} \quad (5)$$

Also, let $\mathbf{a}_{r,m} = [a_{m,1} \ a_{m,2} \ \dots \ a_{m,N}] \in \mathbb{R}^{1 \times N}$, be the m th row of \mathbf{A} . Then at the m th feedback mini-slot, the n th user scales its feedback information by $a_{m,n}$ and feeds back the combination to the relay which simply forwards what it receives to the BS. The relay can feed back the received measurements using FD or HD modes. In the HD case, the relay receives the following at the m th feedback mini-slot

$$y_r^{(m)} = \mathbf{a}_{r,m} \mathbf{x} + z_r^{(m)}, \quad (6)$$

In the case of FD feedback, the BS will have to account for the relay self-interference which does not exist in the case of HD feedback.

$$\begin{aligned} y_r^{(1)} &= \mathbf{a}_{r,1} \mathbf{x} + z_r^{(1)} \\ y_r^{(m)} &= \mathbf{a}_{r,m} \mathbf{x} + \rho y_r^{(m-1)} + z_r^{(m)}, \quad m = 2, 3, \dots, M \end{aligned} \quad (7)$$

$$(8)$$

where $y_r^{(m)}$ is the received signal at the relay at the m th feedback mini-slot, $\rho y_r^{(m-1)}$ is the self-interference residual at time m , and $z_r^{(m)}$ is the relay noise component at the m th mini-slot. For the sake of simplicity, we assume that ρ is small enough such that $|\rho|^k \approx 0, \forall k \geq J$, where J is a constant, so that all factors multiplied by $|\rho|^k$ will be ignored for $k \geq J$. With this in mind, we can transform the recursive expression in (7) and (7) to the following form

$$y_r^{(m)} = \sum_{k=0}^{\min(J-1, m-1)} \rho^k \mathbf{a}_{r, m-k} \mathbf{x} + \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, m-1)} \rho^k z_r^{(m-k)}. \quad (9)$$

The term $\frac{J}{2}$ in the second term of (9) is due to the assumption of the convergence of the noise variance to zero for values of $k \geq \lceil \frac{J}{2} \rceil$, i.e. $\mathbb{E}[|\rho^{\frac{J}{2}} z_r|^2] \approx 0$. After receiving the m th feedback measurement, the relay forwards that signal to the BS. After equalizing by the BS channel, the received m th measurement at the BS becomes

$$y_s^{(m)} = \sum_{k=0}^{\min(J-1, m-1)} \rho^k \mathbf{a}_{r, m-k} \mathbf{x} + \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, m-1)} \rho^k z_r^{(m-k)} + \frac{w^{(m)}}{f}, \quad (10)$$

where $w^{(m)}$ is the noise term at the BS at the m th reception instant.

III. OPPORTUNISTIC USER SELECTION AND FEEDBACK SNR ESTIMATION

In this section, we present the proposed CS-based feedback algorithm and outline the steps required to obtain the identity and the receive SNR of the strong users at the BS. We first demonstrate our approach in the simpler case of HD relaying and then extend it to the case of FD relaying.

A. Feedback Threshold

To apply CS theory, only a few strong users are allowed to feedback. This creates a sparse user regime. This is done by using a feedback threshold⁴. The feedback threshold γ_{th} is optimized to meet a target scheduling outage probability \mathcal{P}_o . The outage probability is defined as the probability that all users fail to report an

SNR above γ_{th} . The scheduling outage probability can be calculated as

$$\mathcal{P}_o = \mathbb{P}(\gamma_n < \gamma_{\text{th}}, \text{ for all } n = 1, \dots, N) = [F_\gamma(\gamma_{\text{th}})]^N \quad (11)$$

where S is the number of users that report an SNR above the threshold and $\bar{S} = \mathbb{E}[S]$ is its mean value, $F_\gamma(\gamma_{\text{th}}) = 1 - \exp(-\frac{\gamma_{\text{th}}}{\bar{\gamma}})$ is the SNR CDF at the users side. From (11), the feedback threshold can be calculated as

$$\gamma_{\text{th}} = F_\gamma^{-1}(\mathcal{P}_o^{1/N}), \quad (12)$$

where \mathcal{P}_o is assumed to be very small.

B. User ID Estimation

For the BS to make a user selection decision, it has to estimate the feedback vector \mathbf{x} . In what follows, we show that we can reliably estimate \mathbf{x} and hence identify a set of strong users. More precisely, we show that in both feedback scenarios (HD and FD) recovering the feedback vector simplifies to solving an underdetermined linear system.

1) *HD feedback*: In the case of HD feedback, the feedback is communicated over orthogonal channels through the two-hop network via the relay. In other words, the relay listens to the users' feedback in one time slot and forwards what it gets in an independent time slot. Thus from (6), we have the following linear system :

$$\mathbf{y}_s = \mathbf{A} \mathbf{x} + \hat{\mathbf{z}}, \quad (13)$$

where $\mathbf{y}_s = [y_s^{(1)} \ y_s^{(2)} \ \dots \ y_s^{(M)}]^t$, and $\hat{\mathbf{z}} = [\frac{w^{(1)}}{f} + z_r^{(1)} \ \frac{w^{(2)}}{f} + z_r^{(2)} \ \dots \ \frac{w^{(M)}}{f} + z_r^{(M)}]^t$. Obviously, the noise vector is uncorrelated with covariance matrix $\Sigma_{\hat{\mathbf{z}}} = \mathbb{E}[\hat{\mathbf{z}}\hat{\mathbf{z}}^t]$. To be able to determine $\Sigma_{\hat{\mathbf{z}}}$, we need to evaluate $\mathbb{E}[\frac{1}{|f|^2}]$. We have⁵

$$\mathbb{E} \left[\frac{1}{|f|^2} \right] = \int_0^\infty \frac{x^{d-2}}{\theta^d \Gamma(d)} e^{-\frac{x}{\theta}} dx = \frac{1}{(d-1)\theta}. \quad (14)$$

Thus, $\Sigma_{\hat{\mathbf{z}}} = \left(N_0 + \frac{N_0}{(d-1)\theta} \right) \mathbf{I}_M$. With this in mind, we can recover the support of \mathbf{x} using regular compressive sensing recovery algorithms such as the least absolute shrinkage and selection operator (LASSO) outlined in [11].

⁴Note that the threshold is the same for both relaying techniques HD and FD.

⁵Note that for $K > 0$, $d = \frac{(K+1)^2}{2K+1} = 1 + \frac{K^2}{2K+1}$, so that $d > 1$.

2) *FD feedback*: In this case, we make use of the structure in the received signal in (10) and then applies the theory of Compressive Sensing to recover \mathbf{x} . Interestingly, the problem can be formulated as a block sparse recovery problem where the unknown vector is a block sparse vector with block size J . We denote by $\mathbf{x}_{(n)} = [x_n, \rho x_n, \dots, \rho^{J-1} x_n]^t$, $n = 1, 2, \dots, N$, and by $\mathbf{a}_{c,n}^j = \left[\underbrace{0, 0, \dots, 0}_j, \mathbf{a}_{c,n}(1 : M - j) \right]^t$.

Then, based on (10), we have the following linear system

$$\mathbf{y}_s = \mathbf{B}\mathbf{v} + \hat{\mathbf{z}}, \quad (15)$$

where $\mathbf{B} = [\mathbf{B}_{(1)}, \mathbf{B}_{(2)}, \dots, \mathbf{B}_{(N)}]$, $\mathbf{v} = [\mathbf{x}_{(1)}^t, \mathbf{x}_{(2)}^t, \dots, \mathbf{x}_{(N)}^t]^t$ and $\mathbf{B}_{(n)} = [\mathbf{a}_{c,n}^0, \mathbf{a}_{c,n}^1, \dots, \mathbf{a}_{c,n}^{J-1}]$, $n = 1, 2, \dots, N$ and $\hat{\mathbf{z}} = [\hat{z}_1 \ \hat{z}_2 \ \dots \ \hat{z}_M]^t$ the noise vector at the BS, where

$$\hat{z}_m = \begin{cases} \frac{w_f^{(1)}}{f} + z_r^{(1)}, & m = 1 \\ \frac{w_f^{(m)}}{f} + z_r^{(m)} + \rho z_r^{(m-1)}, & m = 2, 3, \dots, M. \end{cases} \quad (16)$$

It is not difficult to see that the noise at the BS is correlated with correlation matrix $\Sigma_{\hat{\mathbf{z}}} = \mathbb{E}[\hat{\mathbf{z}}\hat{\mathbf{z}}^t]$. In the case where $J = 3$, we have ⁶

$$\Sigma_{\hat{\mathbf{z}}} = \begin{bmatrix} \alpha_1 & \alpha_3 & 0 & \cdots & 0 \\ \alpha_3 & \alpha_2 & \alpha_3 & \cdots & 0 \\ 0 & \alpha_3 & \ddots & \ddots & \vdots \\ \vdots & \cdots & \ddots & \ddots & \alpha_3 \\ 0 & \cdots & \cdots & \alpha_3 & \alpha_2 \end{bmatrix}, \quad (17)$$

where $\alpha_1 = N_0 \left(1 + \frac{1}{(d-1)\theta}\right)$, $\alpha_2 = N_0 \left(1 + \rho^2 + \frac{1}{(d-1)\theta}\right)$ and $\alpha_3 = \rho N_0$.

As shown in (15), the unknown vector \mathbf{v} is block sparse with block size J and sparsity S . Therefore, it is possible to apply the results in [22] to reliably recover \mathbf{v} . As shown in [22], for reliable CS recovery, the number of measurements should scales as $M = C(JS + S \log JN/S)$. Basically, the BS performs CS recovery (e.g. the CoSaMP algorithm in [22]) on the resultant measurements vector \mathbf{y}_s and estimates the ID of the strongest user. If at least one user is detected, the BS performs a subsequent SNR estimation and refinements to be discussed below, otherwise an outage is declared. In the following subsection, we summarize an important

⁶Our analysis is applicable for any value of J , but for presentation purposes, we restrict the analysis to $J = 3$.

result from CS theory for recovering the sparsity pattern of a block-sparse vector in a noisy setting.

C. Block-Sparse Signal Recovery

The theory of compressed sensing permits efficient acquisition and reconstruction of a sparse signal (through multiplication by an appropriate random matrix) from only few measurements. For a sparse vector with size N and sparsity S , it has been shown that robust signal recovery is possible from $M = \mathcal{O}(S \log N/S)$ [12]. One would expect that additional structure imposed on the unknown sparse signal would substantially reduce M without sacrificing the recovery performance. One imposed structure could be the block-sparsity structure, where the locations of the significant entries in the sparse signal cluster in blocks. In the literature, the concept of recovering a block-sparse vector with reduced number of CS measurements has been studied in [13] and [14]. However robustness guarantees are restricted to either conventional CS sparse signals or recovery with noiseless measurements and do not have exact bounds for the required number of CS measurements. Recently in [22], Baraniuk *et al.* showed that robustness guarantees can be achieved with $M = C(JS + S \log JN/S)$ CS measurements, where C is a positive constant and J is the cluster or the block size, which is a substantial improvement over $M = C(JS \log N/S)$ that would be required by conventional CS recovery algorithms. For more analytical details about block-spare recovery, readers are referred to [13]-[22] and references therein.

D. User SNR Estimation

1) *HD feedback*: After CS recovery, the BS obtains information on the location of the non-zeros in \mathbf{x} or equivalently the support of \mathbf{x} denoted by \mathcal{S} , where $|\mathcal{S}| = S$ is the cardinality of \mathcal{S} . Therefore, the linear system in (13) can be rewritten as

$$\mathbf{y}_s = \mathbf{A}_S \mathbf{x}_S + \hat{\mathbf{z}} \quad (18)$$

We can now apply least squares (LS) to estimate the entries of \mathbf{x}_S as follows

$$\begin{aligned} \hat{\mathbf{x}}_S &= (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \mathbf{A}_S^t \mathbf{y}_s \\ &= \mathbf{x}_S + \mathbf{e}_{LS}, \end{aligned} \quad (19)$$

where \mathbf{e}_{LS} is the estimation error vector. Upon conditioning on \mathbf{A}_S , vector \mathbf{e}_{LS} is Gaussian since it results from a linear operation on Gaussian random variables. The covariance of \mathbf{e}_{LS} is given by $\left[(\mathbf{A}_S^t \mathbf{A}_S)^{-1} \right]$. As M

grow to infinity while S is fixed, $(\mathbf{A}_S^t \mathbf{A}_S)^{-1}$ becomes close to its mean $\frac{M}{M-S-1} N_0 \mathbf{I}_S$ [26] in the sense that :

$$(\mathbf{A}_S^t \mathbf{A}_S)^{-1} - \frac{M}{M-S-1} N_0 \mathbf{I}_S \xrightarrow[M \rightarrow +\infty]{a.s.} 0 \quad (20)$$

The convergence holds in the almost sure sense and shows that the variance of the error is almost the same for all vector entries. In light of this result, we will consider for simplicity that \mathbf{e}_{LS} is Gaussian vector whose entries are independent Gaussian variables with zero-mean and variance $\sigma_{e,LS}^2 = \frac{M}{M-S-1} N_0$.

2) *FD feedback*: Similarly, in the case of FD feedback, the BS recovers the support of \mathbf{v} denoted by \mathcal{I} , where $|\mathcal{I}| = JS$. Therefore, the linear system in (15) can be rewritten as

$$\mathbf{y}_s = \mathbf{B}_{\mathcal{I}} \mathbf{v}_{\mathcal{I}} + \hat{\mathbf{z}}, \quad (21)$$

Since the covariance matrix $\Sigma_{\hat{\mathbf{z}}}$ is correlated, one can apply the Best Linear Unbiased Estimator (BLUE) ⁷ to estimate the entries of $\mathbf{v}_{\mathcal{I}}$ as follows [23]

$$\begin{aligned} \hat{\mathbf{v}}_{\mathcal{I}} &= (\mathbf{B}_{\mathcal{I}}^t \Sigma_{\hat{\mathbf{z}}}^{-1} \mathbf{B}_{\mathcal{I}})^{-1} \mathbf{B}_{\mathcal{I}}^t \Sigma_{\hat{\mathbf{z}}}^{-1} \mathbf{y}_s \\ &= \mathbf{v}_{\mathcal{I}} + \mathbf{e}_{BLUE}, \end{aligned} \quad (22)$$

where \mathbf{e}_{BLUE} is the BLUE estimation error. Similarly to the HD feedback scenario, upon conditioning on $\mathbf{B}_{\mathcal{I}}$, vector \mathbf{e}_{BLUE} is Gaussian with zero-mean and covariance $(\mathbf{B}_{\mathcal{I}}^t \Sigma_{\hat{\mathbf{z}}}^{-1} \mathbf{B}_{\mathcal{I}})^{-1}$. When M tends to infinity with S and J fixed, we can prove that:

Lemma 1: Assume that $M \rightarrow +\infty$ with S and J fixed. Assume that $\limsup_M \|\Sigma_{\hat{\mathbf{z}}}^{-1}\| < +\infty$. Then,

$$(\mathbf{B}_{\mathcal{I}}^t \Sigma_{\hat{\mathbf{z}}}^{-1} \mathbf{B}_{\mathcal{I}})^{-1} - \frac{M \mathbf{I}_{JS}}{\text{tr}(\Sigma_{\hat{\mathbf{z}}}^{-1})} \xrightarrow[M \rightarrow +\infty]{a.s.} 0. \quad (23)$$

Proof: See Appendix A for proof. ■

Again, we note that the above convergence suggests considering the noise at the output of the BLUE estimator as a Gaussian random vector whose entries are independent with zero-mean and variance $\sigma_{e,BLUE}^2$ given by:

$$\sigma_{e,BLUE}^2 = \frac{M}{\text{tr}(\Sigma_{\hat{\mathbf{z}}}^{-1})} \quad (24)$$

E. SNR Back-Off

As stated in (19) and (22), the LS and BLUE estimators are noisy which means that the estimated SNR can be higher or lower than the actual one. This is problematic, since an estimated SNR higher than the actual one results in a transmission rate higher than the maximum rate the end user can support. To deal with

this, we back-off on the estimated noisy SNRs. From (19) and (22), each entry of $\hat{\mathbf{x}}_S$ and $\hat{\mathbf{v}}_{\mathcal{I}}$ respectively can be represented in a scalar equation as $\hat{\gamma} = \gamma + e$, where γ and $\hat{\gamma}$ stands for the actual and the estimated SNRs respectively, and e is the Gaussian error. Our back-off strategy simply subtracts an amount Δ from γ . Hence, $\hat{\gamma}$ becomes

$$\hat{\gamma} = \gamma + e - \Delta \quad (25)$$

To characterize the outage when the estimated SNR is higher than the actual one, we define η as the back-off efficiency which is simply given by

$$\begin{aligned} \eta &= \mathbb{P}(\hat{\gamma} \leq \gamma) \\ &= \mathbb{P}(e \leq \Delta) \\ &= 1 - Q\left(\frac{\Delta}{\sigma_e}\right), \end{aligned} \quad (26)$$

where σ_e is either the LS or the BLUE error variance. Obviously, the performance of the proposed algorithm will depend on Δ . A low value of Δ may result in an outage while a high value will result in a low rate. The optimal value will be derived in the following section.

IV. PERFORMANCE ANALYSIS

The performance of the proposed feedback algorithm will be evaluated based on three performance criteria: i) the feedback load, ii) the achievable rate, and iii) the achievable throughput. Obviously, each of these performance criteria will depend on the feedback relaying mode, i.e. HD or FD feedback. HD feedback results in lower noise variance at the BS at the expense of a larger feedback overhead, while FD feedback results in a higher noise variance at the BS (due to the relay self-interference) but results in a lower feedback overhead.

A. Feedback Load

The feedback load L is defined as the total number of feedback mini-slots required for the BS to make a user scheduling decision. In other words, L is the total number of measurements required to have robust CS recovery at the BS and it is given by

$$L_{\text{HD}} = 2C\bar{S} \log N/\bar{S}$$

in the case of HD feedback⁸, and

$$L_{\text{FD}} = C(J\bar{S} + \bar{S} \log JN/\bar{S}),$$

⁸In HD feedback ($J = 1$), the relay receives and forwards the users' feedback information using two orthogonal channels. Therefore, the BS requires twice the number of mini-slots for feedback reception when compared to FD feedback.

⁷When the noise is white as in eq. (18), LS and BLUE are equivalent.

in the case FD feedback [22]. The term \bar{S} represents the average number of users that feedback and it is given by

$$\begin{aligned}\bar{S} &= \sum_{n=1}^N n \binom{N}{n} (1 - F_\gamma(\gamma_{th}))^n F_\gamma(\gamma_{th})^{N-n} \\ &= N \left(1 - \mathcal{P}_0^{1/N}\right).\end{aligned}\quad (27)$$

B. Achievable Rate

The achievable rate for the HD and FD cases can be expressed as

$$\mathcal{R} = \mathbb{E} \left[\log \left(1 + \gamma_{eq} - \Delta \right) (1 - \mathcal{P}_o) \left(1 - Q \left(\frac{\Delta}{\sigma_e} \right) \right) \right], \quad (28)$$

where

$$\gamma_{eq} = \min \left(\frac{P_s |f|^2}{P_r \rho^2 + N_0}, \gamma_{n^*} \right), \quad (29)$$

where n^* is defined in eq. (2), and $\sigma_e^2 = \sigma_{e,LS}^2$ in the case of HD feedback, $\sigma_e^2 = \sigma_{e,BLUE}^2$ otherwise. Using the Jensen's inequality, (28) can be upper bounded as⁹

$$\mathcal{R} \leq \log \left(1 + \mathbb{E}[\gamma_{eq}] - \Delta \right) (1 - \mathcal{P}_o) \left(1 - Q \left(\frac{\Delta}{\sigma_e} \right) \right). \quad (30)$$

It is clear from (30) that the achievable rate depends on the back-off value Δ . To obtain the optimal value of Δ , we maximize the achievable rate with respect to Δ . To do that, we differentiate the upper bound on \mathcal{R} with respect to Δ and equate the result to zero. We denote by Δ^* , the optimal back-off, then Δ^* satisfies

$$\begin{aligned}&\left(\frac{1 + \bar{\gamma}_{eq} - \Delta^*}{\sqrt{2\pi}\sigma_e} \right) \exp \left(-\frac{(\Delta^*)^2}{2\sigma_e^2} \right) \log \left(1 + \bar{\gamma}_{eq} - \Delta^* \right) \\ &+ Q \left(\frac{\Delta^*}{\sigma_e} \right) = 1\end{aligned}\quad (31)$$

where, $\bar{\gamma}_{eq} = \mathbb{E}[\gamma_{eq}]$. Then, the value of Δ^* that satisfies (31) is used in (28).

It now remains to derive the average end-to-end SNR at the strongest user, i.e. $\mathbb{E}[\gamma_{eq}]$. To achieve this, we first need to derive the PDF of γ_{eq} , $p_{\gamma_{eq}}(x)$. Since γ_{eq} is simply the minimum of two independent random variables, its CDF can be derived as follows

$$\begin{aligned}F_{\gamma_{eq}}(x) &= 1 - \left[1 - F_{|f|^2/(\rho^2+N_0)}(x) \right] \left[1 - F_{\gamma^*}(x) \right] \\ &= 1 - \left[1 - \frac{\gamma(d, \frac{\lambda x}{\theta})}{\Gamma(d)} \right] \left[1 - \left(1 - e^{-\frac{x}{\bar{\gamma}}} \right)^N \right]\end{aligned}\quad (32)$$

⁹Note that the inequality comes also from the fact that we didn't include the CS recovery probability in the rate expression. This probability is not available in closed forms for block-CS recovery.

where $\lambda = \rho^2 + N_0$. Therefore, $p_{\gamma_{eq}}(x)$ is given by

$$\begin{aligned}p_{\gamma_{eq}}(x) &= \frac{\partial F_{\gamma_{eq}}(x)}{\partial x} \\ &= \underbrace{\frac{N}{\bar{\gamma}\Gamma(d)} e^{-\frac{x}{\bar{\gamma}}} \left(1 - e^{-\frac{x}{\bar{\gamma}}} \right)^{N-1} \Gamma \left(d, \frac{\lambda x}{\theta} \right)}_{G_1(x)} \\ &\quad + \underbrace{\left[1 - \left(1 - e^{-\frac{x}{\bar{\gamma}}} \right)^N \right] \frac{1}{\Gamma(d)} \left(\frac{\lambda}{\theta} \right)^d x^{d-1} e^{-\frac{\lambda x}{\theta}}}_{G_2(x)}\end{aligned}\quad (33)$$

Hence,

$$\begin{aligned}\bar{\gamma}_{eq} &= \int_0^\infty x p_{\gamma_{eq}}(x) dx \\ &= \underbrace{\int_0^\infty x G_1(x) dx}_{\mu_1} + \underbrace{\int_0^\infty x G_2(x) dx}_{\mu_2}\end{aligned}\quad (34)$$

To be able to derive the integrals in (34), we use the fact that $\left(1 - e^{-\frac{x}{\bar{\gamma}}} \right)^N = \sum_{n=0}^N \binom{N}{n} (-1)^n e^{-\frac{nx}{\bar{\gamma}}}$ and that $\int_0^\infty x^{\mu-1} e^{-\beta x} \Gamma(\nu, \alpha x) dx = \frac{\alpha^\nu \Gamma(\mu+\nu)}{\mu(\alpha+\beta)^{\mu+\nu}} {}_2F_1 \left(1, \mu + \nu; \mu + 1; \frac{\beta}{\alpha+\beta} \right)$ [28], where ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the Gaussian hypergeometric function. Therefore, the first integral in (34) can be evaluated as

$$\begin{aligned}\mu_1 &= \frac{N\lambda^d}{2\bar{\gamma}d(d+1)\theta^d} \sum_{n=0}^{N-1} \binom{N-1}{n} (-1)^n \left(\frac{\lambda}{\theta} + \frac{n+1}{\bar{\gamma}} \right)^{-d-2} \\ &\quad \times {}_2F_1 \left(1, d+2; 3; \frac{(n+1)/\bar{\gamma}}{\lambda/\theta + (n+1)/\bar{\gamma}} \right).\end{aligned}\quad (35)$$

Similarly, the second integral in (34) can be calculated as

$$\mu_2 = d \left(\frac{\lambda}{\theta} \right)^d \sum_{n=1}^N \binom{N}{n} (-1)^{n+1} \left(\frac{\lambda}{\theta} + \frac{n}{\bar{\gamma}} \right)^{-d-1} \quad (36)$$

Finally,

$$\bar{\gamma}_{eq} = \mu_1 + \mu_2 \quad (37)$$

C. Achievable Throughput

In the previous section, it was assumed that the amount of air-time reserved to feedback is negligible as compared to the transmission time. This assumption does not give much insight for practical scenarios. In this paper, we define the achievable throughput as the number of transmitted bits per unit time (bps/Hz). The network throughput can be explicitly given by

$$\begin{aligned}\mathcal{T} &= \mathcal{R} \frac{(T_c - LT_{ms})}{T_c} \\ &= \mathcal{R} (1 - L\tau)\end{aligned}\quad (38)$$

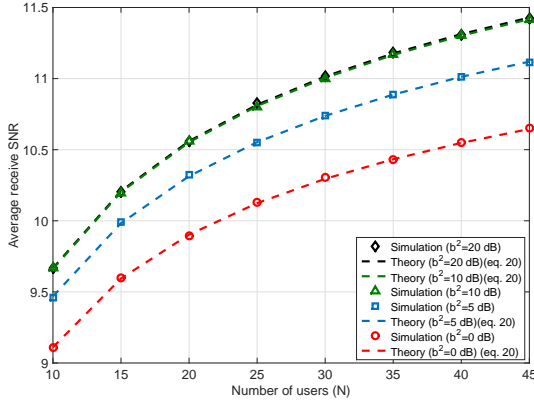


Figure 3. Average Receive SNR versus the number of users (N), $\sigma^2 = 0$ dB, $\sigma_g^2 = 5$ dB, $N_0 = -15$ dB

where T_c is the channel coherence time, T_{ms} is the time needed to transmit one feedback mini-slot and $\tau = \frac{T_{ms}}{T_c}$ is the normalized mini-slot time (see Figure 2).

V. NUMERICAL RESULTS

For all simulations, we set $C = 2$ and we use the CoSaMP algorithm proposed in [22]¹⁰ for CS recovery. Unless otherwise specified, we set: $P_s = P_r = 1$, $b^2 = 20$ dB, $\sigma^2 = 0$ dB, $\sigma_g^2 = 5$ dB, $N_0 = -15$ dB, $\mathcal{P}_0 = 0.01$, and $J = 3$. Throughout this section, we compare the proposed scheme (scheme. 1) with the following schemes:

- Noiseless Dedicated Feedback (scheme. 2): All users feedback in independent noise-free, interference-free feedback channels (via the relay).
- Random User Selection (scheme. 3): For each transmission period, a user is randomly selected and allocated the downlink channel resource.
- Proposed without back-off (scheme. 4): The same as scheme. 1), but without any back-off (SNR refinement) strategy.

In Figure 3, we compare the theoretical result obtained in eq. (37) with the numerical (simulated) average SNR. For all values of the average LOS component b^2 , the theoretical and the numerical results exactly match, which validates the formula in eq. (37). It is also clear from Figure 3 that increasing b^2 increases the average receive SNR as well. However, we notice that the average receive SNR converges for $b^2 \geq 10$ dB, this can be explained by the fact that the receive SNR is constrained by the SNR of the second hop. In Figure 4, it is clear in one hand that scheme. 2 offers the best performance in terms

¹⁰CoSaMP recovery algorithm that we use to recover \mathbf{v} is available at <http://dsp.rice.edu>. For HD feedback, the group size becomes 1.

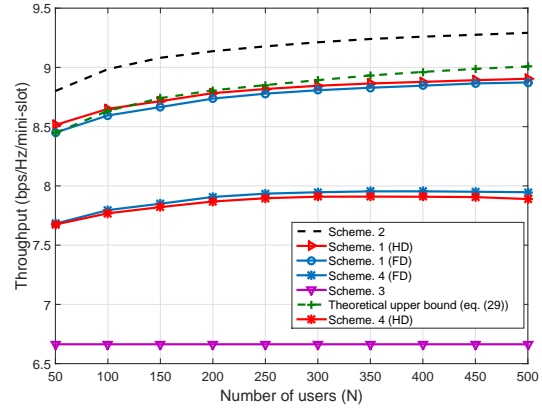


Figure 4. Average Rate versus the number of users (N), $b^2 = 20$ dB, $\sigma^2 = 0$ dB, $\sigma_g^2 = 5$ dB, and $\mathcal{P}_0 = 10^{-2}$.

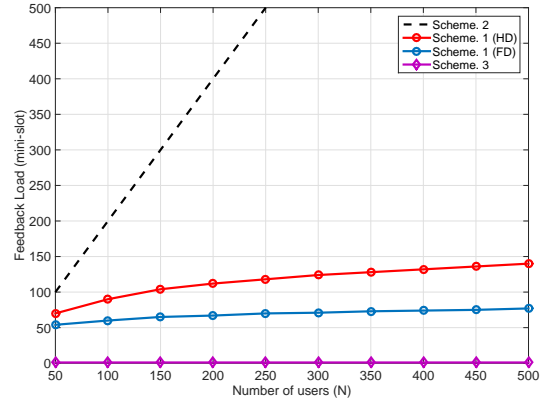


Figure 5. Average Feedback Load versus the number of users (N)

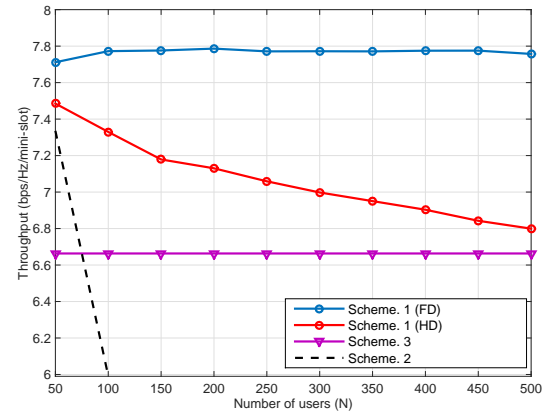


Figure 6. Average Throughput versus the number of users (N), $\tau = 1/600$

of average rate, which is expected since perfect channel state information (CSI) is available at the BS. In the other hand, both versions of the scheme. 1 (HD and

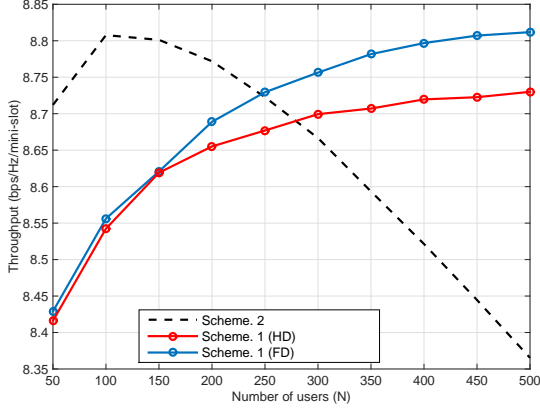


Figure 7. Average Throughput versus the number of users (N), $\tau = 10^{-4}$

FD feedback) offer a comparable performance with a gap of 0.42 bps/Hz as compared to scheme. 2, where the HD version is slightly better than the FD version as it is free from interference. Figure 4 also shows that the performance of the proposed scheme without back-off (scheme. 4) suffers a rate-hit as compared to scheme. 1 and scheme. 2, and hence, the BS has to perform a back-off strategy in order to account for the feedback noise. Random user selection (scheme. 3) is shown to yield the worst performance since no feedback is communicated to the BS.

In Figure 5, we evaluate the performance in terms of feedback load, where we assume that the amount of time needed for channel estimation is negligible compared to the feedback air-time. As shown in Figure 5, scheme. 1 consumes much less feedback as compared to scheme. 2 especially at high number of users N . This is expected, since scheme. 2 requires the feedback to grow linearly with N , whereas scheme. 1 requires the feedback to grow logarithmically with N . The feedback load of scheme. 1 with HD feedback is worse than that of scheme. 1 with FD feedback since feedback is performed over the two hops.

To have more insight, we simulate the network throughput for $\tau = 1/600$ in Figure 6. As shown, scheme. 1 offers the best throughput, and scheme. 2 yields the worst throughput since it requires a large amount of feedback. The throughput of the HD version of scheme. 1 is shown to deteriorate with the number of users. The reason for this is that HD feedback requires more feedback time as the number of users grow when compared to FD feedback. In Figure 7, we decrease τ to be $\tau = 10^{-4}$ (i.e. higher coherence time). We first notice that at low number of users ($N < 250$), scheme. 2 performs better than scheme. 1 (due to the reduced

value of τ) and the HD version of scheme. 1 performs slightly lower than scheme. 1 (FD). At high number of users, scheme. 1 (FD) performs the best. Again this is due to its lower feedback requirements.

VI. CONCLUSION

In this paper, we proposed an opportunistic CS-based feedback algorithm for a relay-aided multiuser network. Instead of allocating a feedback channel for each user, all users are allocated a pool of shared channels for feedback transmissions. We derived the end-to-end receive SNR at the destination user and the post-CS detection/refinement error variance. Moreover, we considered both HD feedback and FD feedback and showed that FD feedback offers comparable rate performance when compared to HD feedback. In addition to this, FD feedback results in lower feedback load, especially at high number of users, and therefore, offers better throughput than HD feedback.

APPENDIX A

PROOF OF LEMMA 1

Let $\mathbf{H} = \mathbf{B}_{\mathcal{I}}^t \boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1} \mathbf{B}_{\mathcal{I}}$. Matrix \mathbf{H} is of finite size JS . It thus suffices to study the convergence of the elements of \mathbf{H} . Using results on the convergence of quadratic forms [27, Lemma5, Lemma4], we can establish that the off-diagonal entries of \mathbf{H} given by

$$(\mathbf{a}_{c,n}^j)^t \boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1} \mathbf{a}_{c,k}^l, \forall (n,j) \neq (k,l)$$

converge almost surely to zero, while the diagonal entries converge almost surely to their mean, i.e.,

$$(\mathbf{a}_{c,n}^j)^t \boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1} \mathbf{a}_{c,n}^j - d_{j,n} \xrightarrow[M \rightarrow +\infty]{a.s.} 0.$$

with $d_{j,n} \triangleq \mathbb{E} \left((\mathbf{a}_{c,n}^j)^t \boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1} \mathbf{a}_{c,n}^j \right)$ reading as:

$$d_{j,n} = \frac{1}{M} \left(\text{tr}(\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1}) - \sum_{k=1}^j [\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1}]_{k,k} \right)$$

In the expression of $d_{j,n}$, the dominant term is $\frac{1}{M} \text{tr}(\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1})$ while the second term converges almost surely to zero, since

$$\frac{1}{M} \sum_{k=1}^j [\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1}]_{k,k} \leq \frac{J \|\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1}\|}{M} \leq \frac{J}{M} \limsup \|\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1}\|$$

We therefore obtain:

$$\mathbf{H}^{-1} - \frac{M \mathbf{I}_{JS}}{\text{tr}(\boldsymbol{\Sigma}_{\hat{\mathbf{z}}}^{-1})} \xrightarrow[M \rightarrow +\infty]{a.s.} 0,$$

In Figure 8, we plot the exact and almost sure limit (lemma 1) of the noise variance versus M/\bar{S} . From the figure we notice that the lower bound on $\sigma_{e,BLUE}^2$ is tight for moderate to high ratio of M/\bar{S} . This concludes the proof of the lemma.

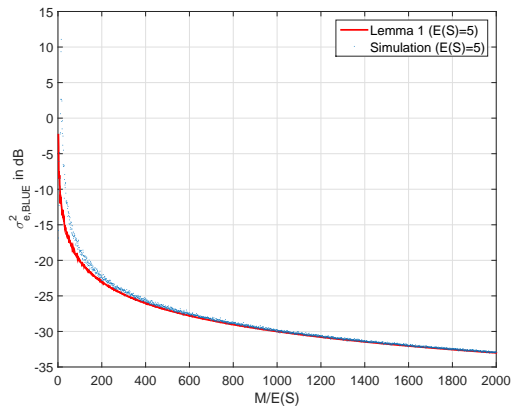


Figure 8. BLUE error variance: Exact versus theoretical limit (Lemma 1).

REFERENCES

- [1] Fan Sun, Tae Min Kim, Paulraj, A.J. and De Carvalho, E. and Popovski, P., "Cell-Edge Multi-User Relaying with Overhearing", *IEEE Communications Letters* vol. 17, no. 6, pp. 1160-1163, June. 2013.
- [2] V. Chandrasekar, J. Andrews, and A. Gatherer, "Femto cell networks: A survey," *IEEE Commun. Mag.*, vol. 46, pp. 59-67, Sept. 2008.
- [3] A. Damnjanovic, J. Montozo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP Heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, pp. 10-21, June 2011.
- [4] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [5] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037-3063, Sept. 2005.
- [6] Y. Jing, and H. Jafarkhani, "Single and multiple relay selection schemes and their achievable diversity orders," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1414-1423, Mar. 2009.
- [7] I. Krikidis, H. Suraweera, and Chau Yuen, "Amplify-and-forward with full-duplex relay selection," *IEEE Tran. Wireless Commun.*, vol. 11, pp. 4381-4393, Dec. 2012.
- [8] B. Rankov and A. Wittneben, "Spectral efficient protocols for half-duplex fading relay channels, *IEEE J. on Selecte Areas in Commun.*, vol. 25, no. 2, pp. 379389, 2007.
- [9] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, Nov. 2006.
- [10] S. Catreux, V. Erceg, D. Gesbert, and R. Heath, "Adaptive Modulation and MIMO Coding for Broadband Wireless Data Networks," *IEEE Commun. Mag.*, vol. 40, pp. 108-115, June 2002.
- [11] E. Candes and Y. Plan, "Near-Ideal Model Selection by l_1 Minimization," *Ann. Statist.*, vol. 37, no. 5A, pp. 2145-2177, 2009.
- [12] A. Fletcher, S. Rangan, and V. Goyal, "Necessary and Sufficient Conditions for Sparsity Pattern Recovery," *IEEE Trans. Inform. Theory*, vol. 55, no. 12, pp. 5758-5772, Dec. 2009.
- [13] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Processing*, vol. 57, no. 8, pp. 3075-3085, Aug. 2009.
- [14] J. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp.572-588, Apr. 2006.
- [15] Y. Eldar, and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302-5316, Nov. 2009.
- [16] S. Bhaskaran, L. Davis, A. Grant, S. Hanly, and P. Tune, "Downlink Scheduling using Compressed Sensing," in *Inf. Theory Workshop*, pp. 201-205, June 2009.
- [17] M. Eltayeb, T. Al-Naffouri, and Hamid Bahrami, "Compressive Sensing for Feedback Reduction in MIMO Broadcast Channels" *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3209-3222, Sept. 2014.
- [18] L. Davis, S. Hanly, and S. Bhaskaran, "Channel Estimation and User Selection in the MIMO Broadcast Channel," *Digital Signal Process. J.*, vol. 21 issue 5, Sept. 2011.
- [19] W. Lu, Y. Liu, and D. Wang, "Efficient Feedback Scheme Based on Compressed Sensing in MIMO Wireless Networks," *Elsevier Computers & Electrical Engineering*, available online Mar. 2012.
- [20] A. Damnjanovic, J. Montojo, W. Yongbin, J. Tingfang, L. Tao, M. Vajapeyam, Y. Taesang, S. Osok, D. Malladi, "A survey on 3GPP heterogeneous networks," *IEEE Wireless Commun.* vol.18, no.3, pp.10,21, June 2011.
- [21] J. G. Andrews et al., "Femtocells: Past, Present, and Future," *IEEE J. Selected Areas of Commun.*, Apr. 2012.
- [22] R. Baraniuk, V. Cevher, M. Duarte and C. Hegde, "Model-Based Compressive Sensing" *IEEE TRANSACTIONS ON INFORMATION THEORY*, VOL. 56, NO. 4, April 2010.
- [23] A. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley & Sons, 2003.
- [24] A. Hazmi, J. Rinne, and M. Renfors, "Cancellation of loop interference with exponential profile using autocorrelation method in OFDM based systems," in Proc. 9th IEEE International Conf. Commun. Syst., Sep. 2004, pp. 140-144.
- [25] H. Hamazumi, K. Imamura, N. Iai, K. Shibuya, and M. Sasaki, "A study of a loop interference canceller for the relay stations in an SFN for digital terrestrial broadcasting," in Proc. IEEE Global Telecommun. Conf., vol. 1, Nov. 2000, pp. 167-171.
- [26] G. Coluccia, A. Roumy, and E. Magli, "Exact Performance Analysis of the Oracle Receiver for Compressed Sensing Reconstruction," in *IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, May 2014, pp. 1005-1009.
- [27] S. Wagner, R. Couillet, M. Debbah and D. T. M. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback" *IEEE TRANSACTIONS ON INFORMATION THEORY*, VOL. 58, NO. 7, July 2012.
- [28] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., Academic Press, 2007.