

On the Feedback Reduction of Multiuser Relay Networks using Compressive Sensing

Khalil Elkhali, Mohammed E. Eltayeb, Abla Kammoun, Tareq Y. Al-Naffouri, and Hamid Reza Bahrami

Abstract—This paper presents a comprehensive performance analysis of full-duplex multiuser relay networks employing opportunistic scheduling with noisy and compressive feedback. Specifically, two feedback techniques based on compressive sensing (CS) theory are introduced and their effect on the system performance is analyzed. The problem of joint user identity and signal-to-noise ratio (SNR) estimation at the base-station is casted as a block sparse signal recovery problem in CS. Using existing CS block recovery algorithms, the identity of the strong users is obtained and their corresponding SNRs are estimated using the best linear unbiased estimator (BLUE). To minimize the effect of feedback noise on the estimated SNRs, a back-off strategy that optimally backs-off on the noisy estimated SNRs is introduced, and the error covariance matrix of the noise after CS recovery is derived. Finally, closed-form expressions for the end-to-end SNRs of the system are derived. Numerical results show that the proposed techniques drastically reduce the feedback air-time and achieve a rate close to that obtained by scheduling techniques that require dedicated error-free feedback from all network users. Key findings of this paper suggest that the choice of half-duplex or full-duplex SNR feedback is dependent on the channel coherence interval, and on low coherence intervals, full-duplex feedback is superior to the interference-free half-duplex feedback.

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

I. INTRODUCTION

THE radio spectrum available for wireless services is extremely scarce, while the demand for data traffic in cellular networks is increasing at an exponential rate [1]. As the link efficiency is approaching its fundamental limits, transmission techniques that better utilize the radio spectrum than existing systems do today are desired. Relay technology has been recently proposed as a solution to provide higher transmission rates with broader network coverage [2].

In fact, multiuser relaying has been recently adopted in a number of industry standards such as IEEE 802.16j mobile multi-hop relaying [3] and IEEE 802.11s mesh networks [4]. Relaying techniques can be mainly classified, based on their forwarding strategy and required processing at the relay terminal, as *decode and forward* (DF) or *amplify and forward* (AF) [3,5]–[10]. In DF relaying, the relay decodes the base-station

(BS) data, prior to re-encoding and transmitting it to the user (or destination), whereas in the AF relaying, the relay amplifies and retransmits the received data without decoding it. Relay transmission for both AF and DF relaying can be performed in *half-duplex* (HD) or *full-duplex* (FD) 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. FD relaying is thus more bandwidth efficient (since it requires one channel use for BS-user transmission) however, it suffers from loop interference due to signal leakage between the relay output and input [10]. While FD relaying has been considered impractical in the past due to the loop interference problems, recent advances on antenna technology and signal processing make it realizable with existing technology (see e.g. [10]–[13]).

Traditionally, the performance of relaying techniques has been widely studied for point-to-point links, i.e., BS-relay-user, and the performance analysis in the multiuser case has received very little attention. The works in [14]–[21] and references therein, investigate key performance measures such as outage probability, average symbol error rate and ergodic capacity of single user relaying networks over various fading channels. In multiuser networks, the end users may experience independent time-variant fading conditions and the desired data rate may not be feasible if the destination user experiences a deep fade. This is mainly due to the fact that the end-to-end signal-to-noise ratio (SNR) is dictated by the weakest relay link, i.e., by the minimum of the BS-relay, relay-user link [22]. Recently the work in [23]–[28] investigated several performance measures in multiuser relay networks. For instance, the work in [23], [24], and [25] derived the outage probability for multiuser relay networks. The cumulative distribution function (CDF) of the highest and lowest end-to-end SNR was derived in [26] and [28] respectively, while diversity analysis was performed in [27].

While these prior works shed light on the performance of multiuser relay networks, they suffer from two key limitations: 1) they assume full channel knowledge at the BS and 2) their analysis is limited to HD relaying. These limitations eventually result in a hit on the achievable rate and throughput. Channel knowledge is usually required by the BS to opportunistically serve the user that experiences the best relative channel conditions. This type of user scheduling is referred to as *opportunistic scheduling* and it was shown to maximize the network performance [29]–[34]. For a small number of users, the amount of feedback (usually performed via the relay) might be negligible, however, for large number of users, the

Khalil Elkhali, 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.elkhali,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.

amount of feedback becomes prohibitively large. This may result in significant transmission delays and performance loss since more air-time is invested in user selection rather than data transmission.

The purpose of this paper is to propose a practical bandwidth efficient relaying system that works in the absence of channel state information (CSI). The paper achieves that by using compressive sensing to feedback the CSI of the strongest users using FD relaying. We also consider HD relaying which is a special case of the FD relaying obtained by setting the FD relay self-interference parameter to zero. Similar to [23], we consider a FD relay network with a single BS, a relay with a line-of-sight (LOS) link to the BS, and multiple users. To reduce the feedback overhead, SNR is only requested from a few strong users. This creates a sparse user regime. We exploit the sparsity in the user regime to pose the feedback problem as a block compressive sensing (CS) recovery problem [35]. Block CS recovery exploits the prior information of the signal block size to better differentiate true signal information from recovery artifacts. This leads to a more robust recovery [36]. We consider both HD and FD feedback. To the best of our knowledge, this is the first paper that proposes a CS-based feedback algorithm for multiuser FD relay-aided networks and accounts for the relay self-interference on the feedback link. In summary, the primary contributions of this paper can be summarized as follows

- 1) We propose a HD and a FD compressive feedback techniques for multiuser relay networks. The novelty of the proposed algorithm lies in the fact that it takes the relay interference into account in the FD case and permits joint SNR feedback and the relay interference estimation at the BS.
- 2) We account for the feedback noise and refine the estimated SNR using existing linear estimation techniques. Moreover, we optimally back-off on the noisy equivalent SNR estimates to reduce the likelihood of an SNR over-estimation error.
- 3) We derive the noise variance induced by the SNR refinement procedure based on results from random matrix theory [37,38].
- 4) We analyze the network performance under the relay self-interference in the downlink. More precisely, we derive the average downlink end-to-end SNR.

The remainder of the paper is organized as follows. In Section II, we introduce the system description. We introduce 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 notation : $F_X(\cdot)$, $F_X^{-1}(\cdot)$, $p_X(\cdot)$ and $\mathbf{E}(X)$ 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 to by lower case bold letters. We use the subscript c and r to refer to a matrix column and row respectively, for example

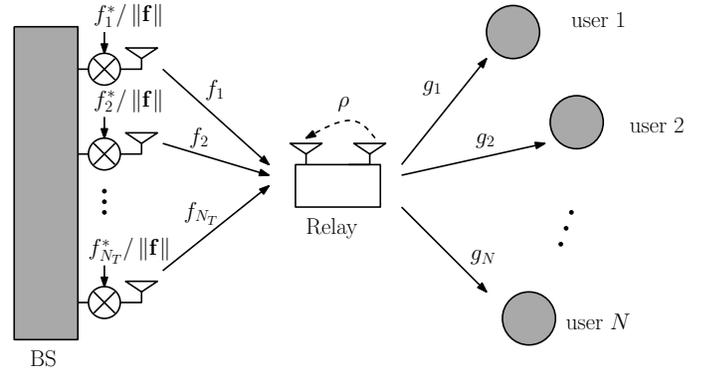


Figure 1: A schematic of a network that consists of a multi-antenna BS, a relay, and multiple users (destinations). Communication between the BS and the users takes place in a two-hop manner via the relay.

$\mathbf{a}_{c,i}$ and $\mathbf{a}_{r,i}$ 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 indexed by the set \mathcal{I} . Moreover, $\text{tr}(\cdot)$ and $[\cdot]^t$ respectively denote the trace and the transpose operators on matrices. We denote by $\|\mathbf{A}\|$ the Frobenius norm of \mathbf{A} and by \otimes the Kronecker product of two matrices. 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 DESCRIPTION

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

A. Downlink Transmission Model

For the downlink, we consider decode and forward relaying as the transmission protocol and a network with N users served by one BS via a single FD relay (no direct path exists between the BS and destination users). As shown in Figure 1, the BS is equipped with N_T antennas, and 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. To model the channels between the BS and the relay, we use the Nakagami- m distribution¹ which covers a wide range of fading distributions encountered in real applications including the Rayleigh fading ($m = 1$). This assumption is in fact realistic since the relay usually exhibits good LOS link with the BS. This is the case where both the relay and the BS are located above the rooftops [23]. The power distribution for Nakagami fading distribution is given by [39]

$$p_{|f|^2}(x) = m^m \frac{x^{m-1}}{\Gamma(m)} \exp(-mx), \quad x \geq 0, \quad (1)$$

¹The channel between the i th BS antenna and the relay, f_i , is modeled as a complex random variable with power distribution following the Gamma distribution

where m is the fading parameter.

Moreover, we model the links between the relay and the users using the Rayleigh distribution. Thus, we assume that the channel from the relay to the n th user ($n = 1, \dots, N$) is complex Gaussian with zero mean and unit variance. Such an assumption can be justified by the fact that mobile users typically suffer severe multipath especially in highly dense and congested areas. Note that this Nakagami-Rayleigh model was also adopted in [23]. The power distribution for the Rayleigh fading distribution is given by

$$p_{|g|^2}(x) = \exp(-x), \quad x \geq 0. \quad (2)$$

To maximize the downlink rate, the destination user is selected based on the following rule

$$n^* = \arg \max_n \gamma_n, \quad (3)$$

where $\gamma_n = \frac{P_r |g_n|^2}{N_0}$ is the instantaneous receive SNR at the n th user, P_r is the relay's transmit power and N_0 is the additive noise power. The relay may apply cancellation techniques as done in [40],[41] to (partly) eliminate the self-interference. Usually this involves estimation of the loop channel 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 [42] and [43], the authors showed that self-interference cancellation can achieve a cancellation of up to 85 dB and 105 dB respectively. As a matter of fact, since typical transmission power for pico cells (or relays) ranges from -6 to 3 dB [1], the effective gain of the loop channel due to the residual interference after cancellation can be modeled as a multiplicative factor $\rho \in [0, 1)$ that remains constant during the transmission interval [10].

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^* , each antenna of the BS broadcasts the symbol $s_{n^*}(t)$ combined with the normalized conjugate of its channel with the relay, i.e., the i th antenna scales its transmission by $\frac{f_i}{\|\mathbf{f}\|}$. Thus, assuming equal power allocation between the BS' antennas, the relay receives

$$y_r(t) = \sqrt{\frac{P_s}{N_T}} \left(\sum_{i=1}^{N_T} \frac{|f_i|^2}{\|\mathbf{f}\|} \right) s_{n^*}(t) + \rho \sqrt{P_r} \hat{s}_{n^*}(t-1) + z_r(t), \quad (4)$$

where $\mathbf{f} = [f_1, f_2, \dots, f_{N_T}]^t$, P_s is the BS total transmission power and the term z_r represents the zero mean variance N_0 additive Gaussian noise at the relay. During the second hop, the relay will decode and forward² the symbol, and the n^* th user receives

$$y_{n^*}(t) = \sqrt{P_r} g_{n^*} \hat{s}_{n^*}(t) + z_{n^*}(t), \quad (5)$$

where $z_{n^*}(t)$ is the zero mean variance N_0 additive Gaussian noise at the n^* th user and the transmitted symbols are assumed to be with unit power.

²Throughout the paper, we only consider decode and forward relaying as the downlink protocol.

B. Feedback Model

We assume a reciprocal TDD setup³ so that the BS-relay and the relay-user channels on the uplink are the same as those of the downlink. 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} ⁴. All channel gains, i.e., $(f_i)_{i=1,2,\dots,N_T}$, ρ 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 a unique signature sequence that represents its identity (ID) 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 (refer to eq. (3)) 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". Thus, let $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_N]^t$ denotes the feedback sparse vector, then x_n is the feedback value of the n th user and it is given by

$$x_n = \begin{cases} \gamma_n & , \gamma_n > \gamma_{th} \\ 0 & , \text{otherwise.} \end{cases} \quad (6)$$

Now, let $\mathbf{a}_{r,t} = [a_{t,1} \ a_{t,2} \ \dots \ a_{t,N}] \in \mathbb{R}^{1 \times N}$, be the t th row of \mathbf{A} . Then at the t th feedback mini-slot, the n th user scales its feedback information by $a_{t,n}$ and feeds back the combination to the relay which simply forwards it to the BS (with unit gain). The relay, using the AF protocol, can feed back the received measurements using FD or HD modes. The reason for using the AF protocol in the feedback phase is due to inevitable feedback collisions at the relay as multiple users feed back simultaneously. Since the relay can not decode feedback collisions, it simply forwards the received signal to the BS, where further processing is carried out to estimate the identity and SNRs of the strong users. In the HD case, the relay receives the following at the t th feedback mini-slot

$$y_r^{(t)} = \mathbf{a}_{r,t} \mathbf{x} + z_r^{(t)}, \quad (7)$$

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

³In this paper we consider a TDD system with reciprocal uplink and downlink channels. The proposed algorithm can also be used to minimize the overhead of frequency-division-duplex systems by allocating a group of shared channels for all users.

⁴One mini-slot refers to the duration needed to communicate one feedback transmission over a single-hop.

HD feedback.

$$y_r^{(1)} = \mathbf{a}_{r,1}\mathbf{x} + z_r^{(1)} \quad (8)$$

$$y_r^{(t)} = \mathbf{a}_{r,t}\mathbf{x} + \rho y_r^{(t-1)} + z_r^{(t)}, t = 2, \dots, M. \quad (9)$$

where $y_r^{(t)}$ is the received signal at the relay at the t th feedback mini-slot, $\rho y_r^{(t-1)}$ is the self-interference residual at time t , and $z_r^{(t)}$ is the relay noise component at the t 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 (8) and (9) to the following truncated form

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

The term $\frac{J}{2}$ in the second term of (10) is due to the assumption of the convergence of the noise variance to zero for noise values with $k \geq \lceil \frac{J}{2} \rceil$, i.e., $\mathbf{E}[|\rho^{\frac{J}{2}} z_r|^2] \approx 0$. After receiving the t th feedback measurement, the relay forwards that signal to the BS. The received t th measurement at the i th BS' antenna after multiplying by $\frac{f_i^*}{\|\mathbf{f}\|}$ is given by

$$y_s^{(t,i)} = \frac{|f_i|^2}{\|\mathbf{f}\|^2} \sum_{k=0}^{\min(J-1, t-1)} \rho^k \mathbf{a}_{r, t-k} \mathbf{x} + \frac{|f_i|^2}{\|\mathbf{f}\|^2} \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, t-1)} \rho^k z_r^{(t-k)} + \frac{f_i^*}{\|\mathbf{f}\|} w^{(t,i)}, \quad (11)$$

where $w^{(t,i)}$ is the noise term at the i th antenna at the t th measurement. Combining the received signals at all the antennas, we get

$$y_s^{(t)} = \left(\sum_{i=1}^{N_T} \frac{|f_i|^2}{\|\mathbf{f}\|^2} \right) \left(\sum_{k=0}^{\min(J-1, t-1)} \rho^k \mathbf{a}_{r, t-k} \mathbf{x} + \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, t-1)} \rho^k z_r^{(t-k)} \right) + \sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|} w^{(t,i)} \\ = \|\mathbf{f}\| \sum_{k=0}^{\min(J-1, t-1)} \rho^k \mathbf{a}_{r, t-k} \mathbf{x} + \|\mathbf{f}\| \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, t-1)} \rho^k z_r^{(t-k)} + \sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|} w^{(t,i)}. \quad (12)$$

Normalizing by $\|\mathbf{f}\|$, we get

$$y_s^{(t)} = \sum_{k=0}^{\min(J-1, t-1)} \rho^k \mathbf{a}_{r, t-k} \mathbf{x} + \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, t-1)} \rho^k z_r^{(t-k)} + \sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|^2} w^{(t,i)}. \quad (13)$$

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, a threshold is applied to allow only a few strong users to feedback. This creates a sparse user regime. The feedback threshold γ_{th} is optimized to meet a target scheduling outage probability \mathcal{P}_0 . 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

$$\mathcal{P}_0 = \mathbb{P}(\gamma_n < \gamma_{\text{th}}, \text{ for all } n = 1, \dots, N) = [F_\gamma(\gamma_{\text{th}})]^N, \quad (14)$$

where $F_\gamma(x) = 1 - \exp\left(-\frac{x}{P_r}\right)$ is the power CDF at the users side. From (14), the feedback threshold can be calculated as

$$\gamma_{\text{th}} = F_\gamma^{-1}(\mathcal{P}_0^{1/N}) = P_r \log\left(\frac{1}{1 - \mathcal{P}_0^{1/N}}\right), \quad (15)$$

where \mathcal{P}_0 is assumed to be very small. Let S be the number of users that report an SNR above the threshold, then S is a Binomial random variable with parameters N and $1 - F_\gamma(\gamma_{\text{th}})$, i.e., $S \sim B(N, 1 - F_\gamma(\gamma_{\text{th}}))$ and its mean value $\bar{S} = \mathbf{E}[S]$ is given by

$$\bar{S} = \sum_{n=1}^N n \binom{N}{n} (1 - F_\gamma(\gamma_{\text{th}}))^n F_\gamma(\gamma_{\text{th}})^{N-n} = N(1 - F_\gamma(\gamma_{\text{th}})) = N(1 - \mathcal{P}_0^{1/N}). \quad (16)$$

B. SNR Feedback and User ID Estimation

For the BS to make a user selection decision, it has to estimate the feedback vector \mathbf{x} (see (6)). 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 under-determined 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. From (13), we have the following linear system

$$\mathbf{y}_s = \mathbf{A}\mathbf{x} + \mathbf{z}, \quad (17)$$

where $\mathbf{y}_s = [y_s^{(1)}, y_s^{(2)}, \dots, y_s^{(M)}]^t$ is the received measurement vector at the BS and $\mathbf{z} = [z_1, z_2, \dots, z_M]^t$ with $z_t = z_r^{(t)} + \sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|^2} w^{(t,i)}$, $t = 1, \dots, M$. Obviously, the noise

vector is uncorrelated with covariance matrix $\Sigma_{\mathbf{z}} = \mathbf{E}[\mathbf{z}\mathbf{z}^t]$. To be able to determine $\Sigma_{\mathbf{z}}$, we need to evaluate the term $\mathbf{E}\left[\left|\sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|^2} w^{(t,i)}\right|^2\right]$. Indeed, we have

$$\begin{aligned} \mathbf{E}\left[\left|\sum_{i=1}^{N_T} \frac{f_i^*}{\|\mathbf{f}\|^2} w^{(t,i)}\right|^2\right] &= \mathbf{E}\left[\left|\sum_{i=1}^{N_T} h_i w^{(t,i)}\right|^2\right] \\ &= \mathbf{E}_{h_i}\left[\sum_{i=1}^{N_T} |h_i|^2 N_0\right] \\ &= N_0 \mathbf{E}_{f_i}\left[\sum_{i=1}^{N_T} \frac{|f_i|^2}{\|\mathbf{f}\|^4}\right] \\ &= N_0 \mathbf{E}_{f_i}\left[\frac{1}{\|\mathbf{f}\|^2}\right], \end{aligned}$$

where $h_i = \frac{f_i^*}{\|\mathbf{f}\|^2}$, $i = 1, \dots, N_T$. Note that $\|\mathbf{f}\|^2$ is the sum of N_T Gamma random variables with shape m and scale parameter $\frac{1}{m}$. Thus, $\|\mathbf{f}\|^2$ is also Gamma distributed with shape $N_T m$ and scale parameter $\frac{1}{m}$. Therefore,

$$\mathbf{E}_{f_i}\left[\frac{1}{\|\mathbf{f}\|^2}\right] = \int_0^\infty (m)^{N_T m} \frac{x^{N_T m - 2}}{\Gamma(N_T m)} \exp(-mx) dx.$$

In the case where $m > \frac{1}{N_T}$ and by a simple change of variables, the above expectation can be easily expressed in closed form using the Gamma function

$$\begin{aligned} \mathbf{E}_{f_i}\left[\frac{1}{\|\mathbf{f}\|^2}\right] &= m \frac{\Gamma(N_T m - 1)}{\Gamma(N_T m)} \\ &= \frac{m}{m N_T - 1}, \end{aligned} \quad (18)$$

where $\frac{\Gamma(m N_T - 1)}{\Gamma(m N_T)} = \frac{1}{m N_T - 1}$. Thus, $\Sigma_{\mathbf{z}}$ is a diagonal matrix with diagonal elements $\frac{N_0 m}{N_T m - 1} + N_0$, i.e., $\Sigma_{\mathbf{z}} = \left(N_0 + \frac{N_0 m}{N_T m - 1}\right) \mathbf{I}_M$. 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 [35].

2) *FD feedback*: As mentioned earlier, in the FD feedback, the received measurements are correlated due to self interference ρ . Thus, we have two ways to process the received signal depending on whether ρ is known or not. In what follows, we discuss each case separately.

a) ρ is known: Based on the expression of the received signal in (11), and since we have perfect knowledge of ρ , the BS can perform successive interference cancellation (SIC) as follows

$$y^{(t)} = \begin{cases} y_s^{(1)}, & t = 1 \\ y_s^{(t)} - \rho y_s^{(t-1)}, & t = 2, 3, \dots, M. \end{cases} \quad (19)$$

Concatenating the obtained measurements yields

$$\mathbf{y} = \begin{bmatrix} y^{(1)} & y^{(2)} & \dots & y^{(M)} \end{bmatrix}^t.$$

It is not difficult to see that the m th entry of \mathbf{y} can be explicitly written as

$$\begin{aligned} y^{(1)} &= \mathbf{a}_{r,1} \mathbf{x} + z_r^{(1)} + \sum_{i=1}^{N_T} h_i w^{(1,i)} \\ y^{(t)} &= \mathbf{a}_{r,t} \mathbf{x} + z_r^{(t)} \\ &\quad + \sum_{i=1}^{N_T} h_i \left(w^{(t,i)} - \rho w^{(t-1,i)} \right), \\ &\quad t = 2, 3, \dots, M. \end{aligned} \quad (20)$$

Based on (20), we have the following linear system

$$\mathbf{y} = \mathbf{A} \mathbf{x} + \mathbf{z}, \quad (21)$$

where $\mathbf{z} = [z_1, z_2, \dots, z_M]^t$ is the effective noise vector at the BS after applying SIC with

$$\begin{aligned} z_1 &= z_r^{(1)} + \sum_{i=1}^{N_T} h_i w^{(1,i)}, t = 1 \\ z_t &= z_r^{(t)} + \sum_{i=1}^{N_T} h_i \left(w^{(t,i)} - \rho w^{(t-1,i)} \right), \\ &\quad t = 2, 3, \dots, M. \end{aligned} \quad (22)$$

Clearly, \mathbf{z} is correlated with covariance matrix $\Sigma_{\mathbf{z}}(\rho)$. The entries of $\Sigma_{\mathbf{z}}(\rho)$ can be obtained as follows

$$\begin{aligned} [\Sigma_{\mathbf{z}}]_{1,1} &= \mathbf{E}\left[\left|z_r^{(1)} + \sum_{i=1}^{N_T} h_i w^{(1,i)}\right|^2\right] \\ &= N_0 \left(1 + \frac{m}{m N_T - 1}\right). \\ [\Sigma_{\mathbf{z}}]_{t,t} &= \mathbf{E}\left[\left|z_r^{(t)} + \sum_{i=1}^{N_T} h_i \left(w^{(t,i)} - \rho w^{(t-1,i)} \right)\right|^2\right] \\ &= N_0 \left(1 + \frac{(1 + \rho^2) m}{m N_T - 1}\right), t = 2, \dots, M. \\ [\Sigma_{\mathbf{z}}]_{i,j} &= \begin{cases} \mathbf{E}[z_i z_j] = -\frac{\rho N_0 m}{m N_T - 1}, & |i - j| = 1, \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

where, we omitted the dependence on ρ for convenience of notation. Based on the linear system in (21), it is possible to apply regular compressive sensing algorithms and recover the support of \mathbf{x} as in the HD case.

b) ρ is unknown: In this case, we make use of the structure in the received signal in (13) and then apply 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 . To this end, define

$$\mathbf{a}_{c,n}^j = \begin{bmatrix} \underbrace{0, 0, \dots, 0}_j, \mathbf{a}_{c,n}(1 : M - j)^t \end{bmatrix}^t.$$

$$\mathbf{B}_{(n)} = [\mathbf{a}_{c,n}^0, \mathbf{a}_{c,n}^1, \dots, \mathbf{a}_{c,n}^{J-1}], n = 1, 2, \dots, N.$$

Then, based on (11), we have the following linear system

$$\mathbf{y}_s = \mathbf{B}\mathbf{v} + \mathbf{z}, \quad (23)$$

where

$$\mathbf{B} = [\mathbf{B}_{(1)}, \mathbf{B}_{(2)}, \dots, \mathbf{B}_{(N)}],$$

$$\mathbf{v} = \mathbf{x} \otimes \begin{bmatrix} 1 \\ \rho \\ \vdots \\ \rho^{J-1} \end{bmatrix},$$

$$\mathbf{z} = [z_1, z_2, \dots, z_M]^t,$$

and \otimes is the Kronecker product. Here, the additive noise \mathbf{z} is defined by

$$z_t = \sum_{k=0}^{\min(\lceil \frac{J}{2} \rceil - 1, t-1)} \rho^k z_r^{(t-k)} + \sum_{i=1}^{N_T} h_i w^{(t,i)},$$

$$t = 1, \dots, M.$$

It is not difficult to see that the noise at the BS is correlated with correlation matrix $\Sigma_{\mathbf{z}}(\rho)$ given by

$$[\Sigma_{\mathbf{z}}]_{1,1} = N_0 \left(1 + \frac{m}{mN_T - 1} \right),$$

$$[\Sigma_{\mathbf{z}}]_{t,t} = N_0 \left(1 + \rho^2 + \frac{m}{mN_T - 1} \right),$$

$$t = 2, \dots, M$$

$$[\Sigma_{\mathbf{z}}]_{i,j} = \begin{cases} \rho^k N_0, & |i-j| = k, 1 \leq k \leq J-1. \\ 0, & \text{otherwise.} \end{cases}$$

As shown in (23), since \mathbf{x} is sparse, the unknown vector \mathbf{v} is block sparse with block size J and sparsity S . Therefore, it is possible to apply the results in [36] to reliably recover \mathbf{v} . As shown in [36], for reliable CS recovery, the number of measurements should scale as $M = C(JS + S \log JN/S)$. Basically, the BS performs CS recovery 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 refinement, otherwise an outage is declared. In the following subsection, we summarize an important 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 compressive 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 $M = \mathcal{O}(S \log N/S)$ [44]. 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 [45] and

[46]. 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 [36], 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 $M = C(JS \log JN/S)$ that would be required by conventional CS recovery algorithms. For more analytical details about block-sparse recovery, the reader is referred to [36,45,47], 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} which corresponds to the identity of the active users. We can now have an improved estimate of the SNR. To this end, note that the system in (17) can be rewritten as

$$\mathbf{y}_s = \mathbf{A}_S \mathbf{x}_S + \mathbf{z}. \quad (24)$$

Since the noise is uncorrelated with covariance $\propto \mathbf{I}_S$, least squares (LS) is the best linear estimator in this case. Upon applying LS, we get

$$\hat{\mathbf{x}}_S = (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \mathbf{A}_S^t \mathbf{y}_s$$

$$= \mathbf{x}_S + \mathbf{e}_0, \quad (25)$$

where \mathbf{e}_0 is the estimation error vector. Upon conditioning on \mathbf{A}_S , the vector \mathbf{e}_0 is Gaussian since it results from a linear operation on Gaussian random variables. The covariance of \mathbf{e}_0 denoted by $\Sigma_{\mathbf{e},0}$ is given by

$$\Sigma_{\mathbf{e},0} = \mathbf{E} [\mathbf{e}_0 \mathbf{e}_0^t]$$

$$= \mathbf{E} \left\{ (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \mathbf{A}_S^t \mathbf{z} \mathbf{z}^t \mathbf{A}_S (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \right\}$$

$$\stackrel{(a)}{=} \mathbf{E} \left\{ (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \mathbf{A}_S^t \Sigma_{\mathbf{z}} \mathbf{A}_S (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \right\}$$

$$= N_0 \left(1 + \frac{m}{mN_T - 1} \right) \mathbf{E} \left\{ (\mathbf{A}_S^t \mathbf{A}_S)^{-1} \right\},$$

where (a) follows from the fact that the entries of \mathbf{z} and \mathbf{A} are mutually independent. Since each row of \mathbf{A}_S is independently drawn from a S -variate normal distribution with zero mean and covariance \mathbf{I}_S , i.e., $\sim \mathcal{N}(\mathbf{0}, \mathbf{I}_S)$. Then, $\mathbf{A}_S^t \mathbf{A}_S \sim \mathcal{W}_S(M, \mathbf{I}_S)$.

Lemma 1: [37] For a central Wishart matrix $\mathbf{W} \sim \mathcal{W}_S(M, \mathbf{I}_S)$ with $M > S$,

$$\mathbf{E} [\text{tr} \{ \mathbf{W}^{-1} \}] = \frac{S}{M - S}. \quad (26)$$

Therefore, the average estimation error is given by

$$\sigma_{\mathbf{e},0}^2 = \frac{1}{S} \mathbf{E} \text{tr} [\mathbf{e}_0 \mathbf{e}_0^t]$$

$$= \frac{N_0}{S} \left(1 + \frac{m}{mN_T - 1} \right) \frac{S}{M - S} \quad (27)$$

$$= \frac{N_0}{M - S} \left(1 + \frac{m}{mN_T - 1} \right).$$

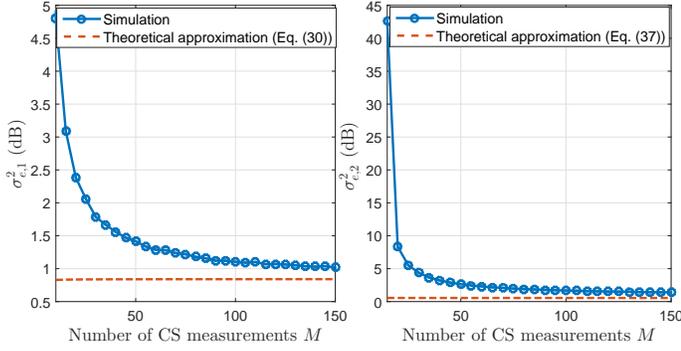


Figure 2: Theoretical approximation and numerical values of the expected noise variance versus the number of measurements M (equations (30) and (37)) with $\rho = 0.5$, $\bar{S} = 5$, $m = 2$, $J = 3$, $N_0 = 1$ and $N_T = 6$.

In light of this result, we will consider for simplicity that \mathbf{e}_0 is Gaussian vector whose entries are independent Gaussian variables with zero-mean and variance $\sigma_{e,0}^2$.

2) *FD Feedback:*

a) ρ is known: Similar to the HD case, after estimating the support \mathcal{S} , the linear system in (21) can be rewritten as

$$\mathbf{y} = \mathbf{A}_S \mathbf{x}_S + \mathbf{z}. \quad (28)$$

To estimate \mathbf{x}_S , we use the best linear unbiased estimator (BLUE) [48] since the noise covariance matrix is not $\propto \mathbf{I}_S$. Upon applying the BLUE, we have

$$\begin{aligned} \hat{\mathbf{x}}_S &= (\mathbf{A}_S^t \Sigma_{\mathbf{z}}^{-1} \mathbf{A}_S)^{-1} \Sigma_{\mathbf{z}}^{-1} \mathbf{A}_S^t \mathbf{y}_S \\ &= \mathbf{x}_S + \mathbf{e}_1, \end{aligned} \quad (29)$$

Following the same approach as in Section III.D. 1, the covariance matrix of \mathbf{e}_1 is given by

$$\begin{aligned} \Sigma_{e,1} &= \mathbf{E} [\mathbf{e}_1 \mathbf{e}_1^t] \\ &= (\mathbf{A}_S^t \Sigma_{\mathbf{z}}^{-1} \mathbf{A}_S)^{-1}. \end{aligned}$$

Unlike the HD case, $\sigma_{e,1}^2 = \frac{1}{\bar{S}} \mathbf{E} \text{tr} [\Sigma_{e,1}]$ is difficult to derive in general. To overcome this issue, we derive an approximate value for $\sigma_{e,1}^2$. This is the subject of the following lemma whose proof is available in Appendix A.

Lemma 2: For large M , the following convergence holds

$$\text{tr} (\mathbf{A}_S^t \Sigma_{\mathbf{z}}^{-1} \mathbf{A}_S)^{-1} - \frac{MS}{\text{tr} (\Sigma_{\mathbf{z}}^{-1})} \xrightarrow{M \rightarrow \infty} 0$$

almost surely.

Proof: See Appendix A for proof. ■

Based on this lemma, for large M we have

$$\sigma_{e,1}^2 \approx \frac{M}{\text{tr} (\Sigma_{\mathbf{z}}^{-1})}. \quad (30)$$

b) ρ is unknown: In this case, the BS recovers the support of \mathbf{v} denoted by \mathcal{I} , where $|\mathcal{I}| = JS$. Therefore, the linear system in (23) can be rewritten as

$$\mathbf{y}_S = \mathbf{B}_{\mathcal{I}} \mathbf{v}_{\mathcal{I}} + \mathbf{z}, \quad (31)$$

Similar to the previous case, to estimate the entries of $\mathbf{v}_{\mathcal{I}}$, we apply the BLUE as follows

$$\begin{aligned} \hat{\mathbf{v}}_{\mathcal{I}} &= (\mathbf{B}_{\mathcal{I}}^t \Sigma_{\mathbf{z}}^{-1} \mathbf{B}_{\mathcal{I}})^{-1} \mathbf{B}_{\mathcal{I}}^t \Sigma_{\mathbf{z}}^{-1} \mathbf{y}_S \\ &= \mathbf{v}_{\mathcal{I}} + \mathbf{e}_2, \end{aligned} \quad (32)$$

where \mathbf{e}_2 is the BLUE estimation error. Now, recall that

$$\hat{\mathbf{v}}_{\mathcal{I}} = \hat{\mathbf{x}}_S \otimes \begin{bmatrix} 1 \\ \rho \\ \vdots \\ \rho^{J-1} \end{bmatrix}.$$

Using this relationship between $\hat{\mathbf{x}}_S$ and $\hat{\mathbf{v}}_{\mathcal{I}}$, we can write

$$\begin{aligned} \hat{\rho} &= \left[\frac{\hat{\mathbf{v}}_{\mathcal{I}} ((i-1)J + j + 1)}{\hat{\mathbf{x}}_S(i)} \right]^{1/j}, \quad j = 1, \dots, J-1, \\ & \quad i = 1, \dots, S. \end{aligned}$$

The last equation shows that we have $(J-1)S$ estimates of ρ which we can average to get a more accurate estimate, as follows

$$\hat{\rho} = \frac{1}{(J-1)S} \sum_{i=1}^S \sum_{j=1}^{J-1} \left[\frac{\hat{\mathbf{v}}_{\mathcal{I}} ((i-1)J + j + 1)}{\hat{\mathbf{x}}_S(i)} \right]^{1/j}, \quad (33)$$

where the first sum is taken over the blocks and the second over the observations in the same block. In the following lemma, we derive the asymptotic equivalent of the BLUE error covariance matrix when ρ is fixed.

Lemma 3: For a fixed ρ , assume that $M \rightarrow +\infty$ with S and J fixed. Assume that $\limsup_M \|\Sigma_{\mathbf{z}}^{-1}(\rho)\| < +\infty$. Then,

$$(\mathbf{B}_{\mathcal{I}}^t \Sigma_{\mathbf{z}}^{-1}(\rho) \mathbf{B}_{\mathcal{I}})^{-1} - \frac{M \mathbf{I}_{JS}}{\text{tr} (\Sigma_{\mathbf{z}}^{-1}(\rho))} \xrightarrow[M \rightarrow +\infty]{a.s.} 0. \quad (34)$$

Proof: See Appendix B for proof. ■

Based on the previous lemma, we can approximate $\Sigma_{e,2}$ as

$$\begin{aligned} \Sigma_{e,2} &= \mathbf{E} [\mathbf{e}_2 \mathbf{e}_2^t] \\ &= (\mathbf{B}_{\mathcal{I}}^t \Sigma_{\mathbf{z}}^{-1}(\rho) \mathbf{B}_{\mathcal{I}})^{-1} \\ &\approx \frac{M \mathbf{I}_{JS}}{\text{tr} (\Sigma_{\mathbf{z}}^{-1}(\rho))}. \end{aligned} \quad (35)$$

Intuitively, an estimate of the BLUE error covariance matrix would be the asymptotic limit derived in (35) with ρ being replaced by $\hat{\rho}$ obtained in (33). In other words,

$$\hat{\Sigma}_{e,2} = \frac{M \mathbf{I}_{JS}}{\text{tr} (\Sigma_{\mathbf{z}}^{-1}(\hat{\rho}))}. \quad (36)$$

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 $\sigma_{e,2}^2$ given by

$$\hat{\sigma}_{e,2}^2 = \frac{M}{\text{tr} (\Sigma_{\mathbf{z}}^{-1}(\hat{\rho}))}. \quad (37)$$

We estimate $\mathbf{v}_{\mathcal{I}}$ and ρ in an iterative way. Starting with an initial estimate of $\mathbf{v}_{\mathcal{I}}^5$, we estimate ρ using (33), then update $\mathbf{v}_{\mathcal{I}}$ using (32).

In Figure 2, the average estimation error for different values of CS measurements M is plotted. In the first subplot, we notice that as M increases, the discrepancy between the true values and the approximation in (30) decreases. The convergence is a bit slower for the second subplot showing the discrepancy between the true value and the approximation in (37). This can be explained by two reasons: 1) the measurement matrix in the case where ρ is unknown is correlated which makes the convergence slower, 2) more measurements are needed to estimate ρ .

E. SNR Back-Off

As stated in (25), (29) and (32), 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 (25), (29) and (32), each entry of $\hat{\mathbf{x}}_{\mathcal{S}}$ 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

$$\hat{\gamma} = \gamma + e - \Delta. \quad (38)$$

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

$$\begin{aligned} \eta &= \mathbb{P}(\hat{\gamma} \leq \gamma) \\ &= \mathbb{P}(e \leq \Delta) \\ &= 1 - Q\left(\frac{\Delta}{\sigma_e}\right), \end{aligned} \quad (39)$$

where $\sigma_e^2 = \sigma_{e,k}^2$, $k = 0, 1, 2$, depending on the scenario as stated earlier. 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 as part of the performance analysis 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 yields a lower feedback overhead.

⁵We can estimate $\mathbf{v}_{\mathcal{I}}$ initially using Least Squares (LS) since LS does not need any statistics.

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.

1) *HD Feedback*: In HD feedback, the relay receives and forwards the users' feedback information using two orthogonal channels. Thus,

$$L_{HD} = 2M, \quad (40)$$

where $M > \lceil J\bar{S} + \bar{S} \log(JN/\bar{S}) \rceil$ [36] is the number of CS measurements. Substituting the value of M in (40), for a block size of $J = 1$, the feedback load becomes

$$L_{HD} = 2C_1\bar{S} \lceil 1 + \log(N/\bar{S}) \rceil, \quad (41)$$

where C_1 is a compressive sensing constant and \bar{S} is the average number of users that feedback.

2) *FD Feedback (ρ is known)*: In this case, we need half the number of measurements needed for the HD case since, each feedback measurement is received and forwarded simultaneously, i.e.,

$$\begin{aligned} L_{FD,1} &= M \\ &= C_1\bar{S} \lceil 1 + \log(N/\bar{S}) \rceil. \end{aligned}$$

3) *FD Feedback (ρ is unknown)*: In this case, the feedback load is equal to the number of measurements needed to have robust block sparse recovery, i.e.,

$$\begin{aligned} L_{FD,2} &= M \\ &= C_2 \lceil J\bar{S} + \bar{S} \log(JN/\bar{S}) \rceil. \end{aligned}$$

One can increase J to approach the actual interference present at the relay, however this comes at the expense of more feedback load. In fact, increasing the block size by 1, will increase the feedback load by $C_2(\bar{S} + \bar{S} \log \frac{J+1}{J}) \approx \mathcal{O}(\bar{S})$.

B. Achievable Rate

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

$$\mathcal{R} = \mathbf{E}[\log(1 + \gamma_{eq} - \Delta)] (1 - \mathcal{P}_o) \left(1 - Q\left(\frac{\Delta}{\sigma_e}\right) \right), \quad (42)$$

where

$$\gamma_{eq} = \min\left(\frac{P_s}{N_T} \sum_{i=1}^{N_t} |f_i|^2, \gamma_{n^*}\right), \quad (43)$$

where n^* is defined in (3). Note that (43) is obtained by taking into account that the downlink protocol is DF. Using the Jensen's inequality, (42) can be tightly upper bounded for high SNR as ⁶

$$\mathcal{R} \lesssim \log(1 + \mathbf{E}[\gamma_{eq}] - \Delta) (1 - \mathcal{P}_o) \left(1 - Q\left(\frac{\Delta}{\sigma_e}\right) \right), \quad (44)$$

⁶Note 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 form for block-CS recovery.

It is clear from (44) that the 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

$$\left(\frac{1 + \bar{\gamma}_{eq} - \Delta^*}{\sqrt{2\pi}\sigma_e}\right) \exp\left(-\frac{(\Delta^*)^2}{2\sigma_e^2}\right) \log(1 + \bar{\gamma}_{eq} - \Delta^*) + Q\left(\frac{\Delta^*}{\sigma_e}\right) = 1, \quad (45)$$

where, $\bar{\gamma}_{eq} = \mathbf{E}[\gamma_{eq}]$. Then, the value of Δ^* that satisfies (45) is used in (42). It now remains to derive the average end-to-end SNR at the strongest user, i.e., $\mathbf{E}[\gamma_{eq}]$. To do so, we first rewrite $\mathbf{E}[\gamma_{eq}]$ using the conditional expectation on ρ as follows

$$\mathbf{E}[\gamma_{eq}] = \mathbf{E}_\rho[\gamma_{eq}|\rho].$$

We start by deriving conditional PDF $p_{\gamma_{eq}|\rho}(x) = p_{\gamma_{eq}(\rho)}(x)$. Since $\gamma_{eq}|\rho$ is simply the minimum of two independent random variables, its CDF can be simply derived as follows

$$F_{\gamma_{eq}(\rho)}(x) = 1 - \left[1 - \frac{\gamma\left(m, \frac{x}{\theta}\right)}{\Gamma(m)}\right] \left[1 - \left(1 - e^{-\frac{x}{\bar{\gamma}}}\right)^N\right], \quad (46)$$

where $\theta = \frac{P_S}{m(P_r\rho^2 + N_0)}$ and $\bar{\gamma} = \frac{P_r}{N_0}$. Therefore, $p_{\gamma_{eq}}(x)$ is given by

$$\begin{aligned} & p_{\gamma_{eq}(\rho)}(x) \\ &= \frac{\partial F_{\gamma_{eq}}(x)}{\partial x} \\ &= \frac{N}{\bar{\gamma}\Gamma(m)} e^{-\frac{x}{\bar{\gamma}}} \left(1 - e^{-\frac{x}{\bar{\gamma}}}\right)^{N-1} \Gamma\left(m, \frac{x}{\theta}\right) \\ & \quad + \underbrace{\left[1 - \left(1 - e^{-\frac{x}{\bar{\gamma}}}\right)^N\right]}_{G_1(x)} \underbrace{\frac{1}{\Gamma(m)\theta^m} x^{m-1} e^{-\frac{x}{\theta}}}_{G_2(x)}. \end{aligned} \quad (47)$$

Hence,

$$\begin{aligned} \bar{\gamma}_{eq}(\rho) &= \mathbf{E}[\gamma_{eq}|\rho] \\ &= \int_0^\infty x p_{\gamma_{eq}(\rho)}(x) dx \\ &= \underbrace{\int_0^\infty x G_1(x) dx}_{\mu_1(\rho)} + \underbrace{\int_0^\infty x G_2(x) dx}_{\mu_2(\rho)}. \end{aligned} \quad (48)$$

To be able to derive the integrals in (48), 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

$$\begin{aligned} & \int_0^\infty x^{\mu-1} e^{-\beta x} \Gamma(\nu, \alpha x) dx \\ &= \frac{\alpha^\nu \Gamma(\mu + \nu)}{\mu(\alpha + \beta)^{\mu + \nu}} {}_2F_1\left(1, \mu + \nu; \mu + 1; \frac{\beta}{\alpha + \beta}\right), \end{aligned}$$

where ${}_2F_1(\cdot, \cdot; \cdot; \cdot)$ is the Gaussian hypergeometric function [49, eq. (6.455)]. Therefore, the first integral in (48) can be

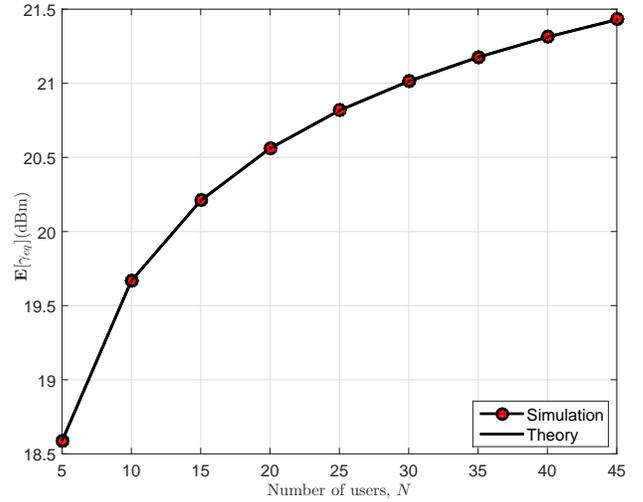


Figure 3: A comparison between the theoretical (50) and simulated average end-to-end SNR versus the number of users N , where $\rho \sim U_{[0,1]}$. The simulation parameters are summarized in Table I.

evaluated as

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

Similarly, the second integral in (48) can be calculated as

$$\begin{aligned} \mu_2(\rho) &= \frac{m}{\theta^m} \sum_{n=1}^N \binom{N}{n} (-1)^{n+1} \\ & \quad \times \left(\frac{1}{\theta} + \frac{n}{\bar{\gamma}}\right)^{-m-1}. \end{aligned}$$

Thus,

$$\bar{\gamma}_{eq}(\rho) = \mu_1(\rho) + \mu_2(\rho). \quad (49)$$

Finally,

$$\mathbf{E}[\gamma_{eq}] = \int_0^1 \bar{\gamma}_{eq}(\xi) p_\rho(\xi) d\xi, \quad (50)$$

where $p_\rho(\xi)$ is the PDF of ρ .

In Figure 3, we validate the result obtained in (50) and compare the theoretical result in (50) with the simulated end-to-end-SNR. The figure shows that the maximum theoretical end-to-end SNR coincide with the simulated maximum end-to-end SNR, and hence validates our theoretical result in (50). The simulation parameters are tabulated in Table I.

C. 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

Parameter	Value	Parameter	Value
P_s	43 dBm	C_1, C_2	2
P_r	15 dBm	N_0	1
N_T	6 antennas	\mathcal{P}_0	0.01
m	2	J	{3, 4, 5}

Table I: Simulation parameters

(bps/Hz). The network throughput can be explicitly given by

$$\begin{aligned} \mathcal{T} &= \mathcal{R} \frac{(T_c - LT_{ms})}{T_c} \\ &= \mathcal{R} (1 - L\tau). \end{aligned} \quad (51)$$

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.

V. NUMERICAL RESULTS AND DISCUSSION

In this section, we study the performance of the proposed CS-based feedback algorithms assuming that ρ follows a uniform distribution over $[0, 1)$, i.e., $\rho \sim U_{[0,1)}$ ⁷. Throughout this section, we use the following abbreviations:

- HD feedback algorithm: (HD).
- FD feedback algorithm with ρ known: (FD-SIC).
- FD feedback algorithm with ρ unknown: (FD-block-CS).

For CS recovery, we use CoSaMP for both conventional ($J = 1$) and block CS recovery which is available from <http://dsp.rice.edu>. We tabulate the simulation parameters as shown in Table I and briefly elaborate on the choice of each parameter. The choice of P_s and P_r is based on the setting released by Qualcomm [50]. The choice of N_T is widely used in the literature [51]. The choice of the shape parameter m is not restricted to $m = 2$, we can choose any value greater than 1 since the BS exhibits good line of sight link with the relay. C_1 and C_2 are compressive sensing constants that are chosen to have good recovery performance. The choice of \mathcal{P}_0 controls the sparsity of the feedback vector. The choice made in Table I ensures that few users feedback their SNRs with small outage probability. In our simulations, we compare the proposed feedback algorithms with the *full feedback* algorithm that requires noiseless and interference-free (i.e., HD feedback) dedicated feedback from all users. Due to its dedicated feedback, the full feedback algorithm requires the largest amount feedback and thus, it is expected to achieve the best rate performance.

In Figure 4, we plot the achievable rate versus the number of users. The figure shows that the rate achieved by proposed FD-SIC matches the rate of the HD feedback algorithm. This can be explained by the fact that both algorithms have the same measurement matrix that is uncorrelated unlike the FD-block-CS algorithm. As a matter of fact, the FD-block-CS has a small gap in the rate performance as compared to the two other algorithms due to the estimation error of ρ and the structure of the measurement matrix. In the same figure, we see that increasing the block-size permits the BS to obtain better estimates of the relay self-interference ρ , (in addition to

⁷The uniform distribution is used in this paper for the sake of simplicity, however any other distribution can be used to model the self interference.

the IDs of the strong users and their relay-user instantaneous SNRs). This results in a better end-to-end (BS-relay-user) channel estimates and improves the achievable rate. We stress the fact that the proposed FD-block-CS is more practical since it assumes no knowledge of ρ and jointly estimates the feedback information and the self interference with slightly more feedback overhead as shown in Figure 5.

In Figure 5, two observations can be concluded. First, FD-block-CS requires slightly additional feedback compared to the FD-SIC in order to estimate the interference however its feedback load is less than the HD since the latter requires the feedback to be communicated in orthogonal channels. Second, the feedback load is the highest for the full feedback algorithm since it requires dedicated noiseless and interference-free feedback from all users.

In Figure 6, we plot the rate performance versus the number of CS measurements M . As shown in Figure 6, as the number of CS measurements increases, the rate of the proposed algorithms increases and converges to a rate that is close to the rate promised by the full feedback algorithm. For the FD-block-CS, the convergence is slower than that of the other two algorithms. Note that we only considered the number of measurements in this figure, not the feedback load. For the HD algorithm to achieve the performance in Figure 6, it needs twice the number of measurements, i.e., $2M$. Thus, for practical scenarios, the HD algorithm is expected to yield a performance degradation due to the large feedback air time that it needs. To show this relation between the feedback air time and the actual rate, we plot the throughput as this metric gives more practical insights on the performance of the different feedback algorithms. We distinguish between two cases: 1) the feedback mini-slot is relatively small as compared to the coherence time ($\tau = 1/3000$, large coherence time) and 2) $\tau = 1/200$ (small coherence time). The values of τ are chosen for illustrative purposes only, however, any values of τ can be taken.

Figures 7 and 8 account for the feedback air-time when considering low and high coherence intervals. If we normalize the feedback mini-slot time to $T_{MS} = 1$ time unit and the whole frame (or coherence interval) to $T_F = 200$ time units, i.e., $\tau = 1/200$ (short coherence interval), Figure 7 shows that the throughput of the full feedback algorithm deteriorates with the number of users. The reason for this is that as the number of users increases, the coherence interval is dominated by feedback traffic rather than data traffic. This leaves little time for data transmission. The proposed algorithm achieves the best throughput since it requires minimum time to capture most of the multiuser diversity without explicit feedback from all users. In Figure 8, we plot the throughput when $\tau = 1/3000$, i.e., when we assume a large coherence interval. For low number of users, the throughput of the full feedback algorithm is shown to be the highest when compared to the reference algorithms. However, the figure shows that the throughput deteriorates after some values of N , i.e., for large number of users, since the amount of feedback traffic increases linearly with the number of users in the full feedback algorithm.

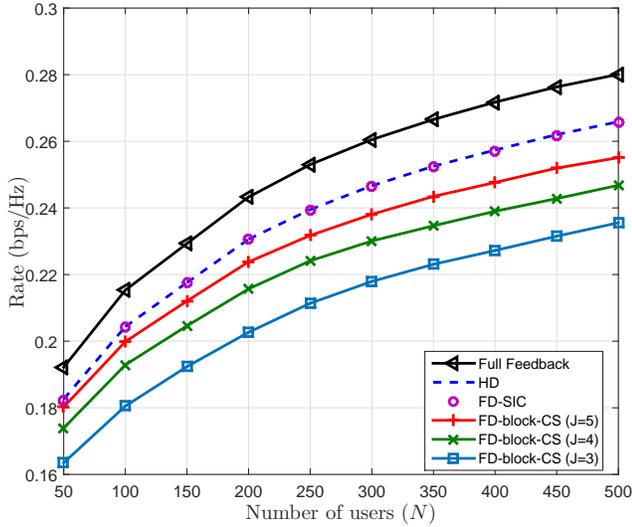


Figure 4: Comparison of the achievable rate versus the number of users N for different feedback techniques. The feedback block-size J is varied from 3-5 when the self-interference ρ is unknown at the relay.

In Figure 9, we plot the mean square error ($\mathbf{E}|\hat{\rho} - \rho|^2$) (MSE) of our proposed interference estimation algorithm highlighted in (33). As shown in Figure 9 and for $N = 300$ users, the estimation is improved by increasing the number of CS measurements, and an approximate MSE of -20 dB is achieved with only 80 CS measurements for $N = 300$ users.

VI. CONCLUSION

In this paper, we investigated the performance of multiuser relaying networks with compressive and noisy SNR feedback. Instead of allocating a feedback channel for each user, all users are allocated a pool of shared channels for feedback transmission. Joint user ID and SNR estimation is performed at the BS using existing compressive sensing recovery algorithms. We derived the end-to-end SNR and the error covariance matrix of the noise after CS recovery which permits fast and efficient off-line performance evaluation. Moreover, we considered both HD feedback and FD feedback and showed that FD feedback with perfect interference knowledge at the BS yields the same rate performance when compared to HD feedback with lower feedback overhead. When the relay interference is not known, the proposed FD feedback technique jointly estimates the feedback SNR, user IDs, and the relay interference, and results in a rate comparable to the HD case with lower feedback load. The findings of the paper suggest the use of full-duplex feedback when the relay interference power is known at the BS and half-duplex feedback when the relay interference power is not known. Moreover, the findings reveal that full-duplex feedback is superior to the interference-free half-duplex feedback in systems with low channel coherence intervals.

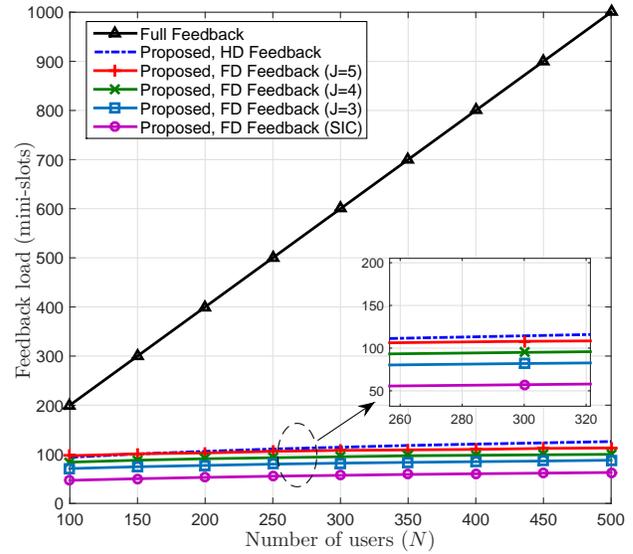


Figure 5: Comparison of the feedback load versus the number of users N for different feedback techniques.

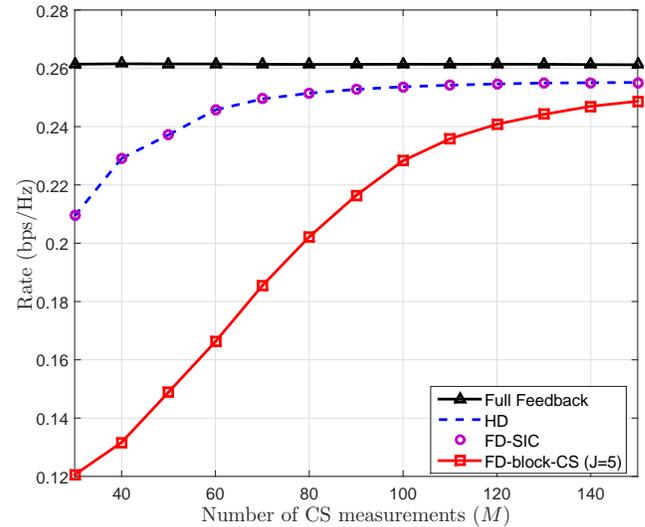


Figure 6: Achievable rate versus the number of CS measurements M for different feedback techniques. The number of users is $N = 300$ users and the block size is $J = 5$.

APPENDIX A

PROOF OF LEMMA 2

Prior to proving the lemma, we provide some useful results from random matrix theory.

Lemma 4: [38] Let $\mathbf{A}_1, \mathbf{A}_2, \dots$, with $\mathbf{A}_M \in \mathbb{C}^{M \times M}$ be a series of random matrices generated by the probability space (Ω, \mathcal{F}, P) such that, for $\omega \in A \subset \Omega$, with $P(A) = 1$, $\|\mathbf{A}_M(\omega)\| < K(\omega) < \infty$, uniformly on M . Let $\mathbf{x}_1, \mathbf{x}_2, \dots$, with $\mathbf{x}_M \in \mathbb{C}^N$, be random vectors of i.i.d entries with zero mean, variance $1/M$, and eighth-order moment of order $\mathcal{O}(1/M^4)$, independent of \mathbf{A}_M . Then

$$\mathbf{x}_M^H \mathbf{A}_M \mathbf{x}_M - \frac{1}{M} \text{tr} \mathbf{A}_M \xrightarrow{M \rightarrow \infty} 0,$$

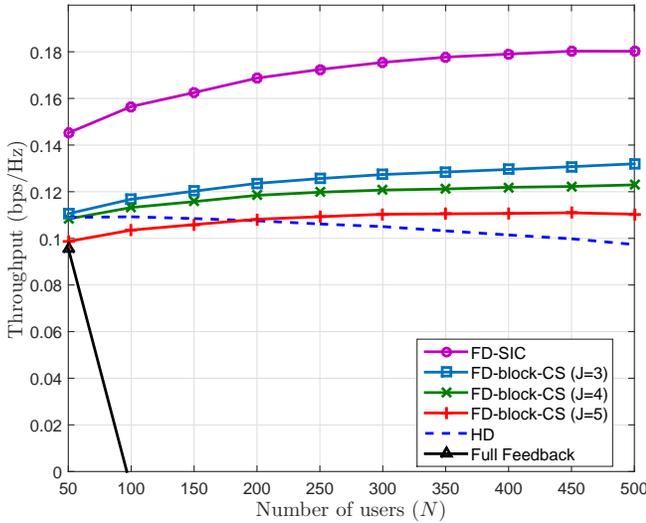


Figure 7: Achievable throughput versus the number of users N for different feedback techniques. The normalized feedback mini-slot is set to $\tau = 1/200$ (small coherence interval).

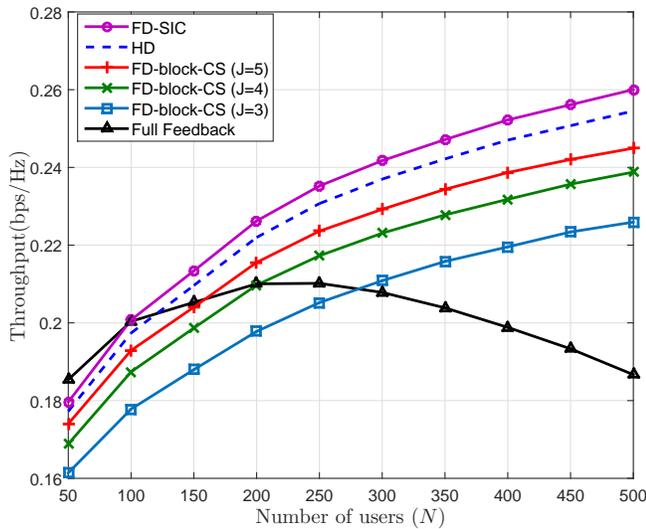


Figure 8: Achievable throughput versus the number of users N for different feedback techniques. The normalized feedback mini-slot is set to $\tau = 1/3000$ (large coherence interval).

and

$$\mathbf{y}_M^H \mathbf{A}_M \mathbf{x}_M \xrightarrow{M \rightarrow \infty} 0$$

almost surely, for $\mathbf{y}_M \in \mathbb{C}^N$ independent of \mathbf{x}_M with standard i.i.d. entries of zero mean and variance $1/M$.

Based on the previous lemma, we can state the following result

$$(\mathbf{A}_S^t \boldsymbol{\Sigma}_z^{-1} \mathbf{A}_S) - \frac{\text{tr}(\boldsymbol{\Sigma}_z^{-1})}{M} \mathbf{I}_S \xrightarrow{M \rightarrow \infty} 0.$$

Consequently,

$$(\mathbf{A}_S^t \boldsymbol{\Sigma}_z^{-1} \mathbf{A}_S)^{-1} - \frac{M}{\text{tr}(\boldsymbol{\Sigma}_z^{-1})} \mathbf{I}_S \xrightarrow{M \rightarrow \infty} 0,$$

$$\text{and } \text{tr}(\mathbf{A}_S^t \boldsymbol{\Sigma}_z^{-1} \mathbf{A}_S)^{-1} - \frac{MS}{\text{tr}(\boldsymbol{\Sigma}_z^{-1})} \xrightarrow{M \rightarrow \infty} 0.$$

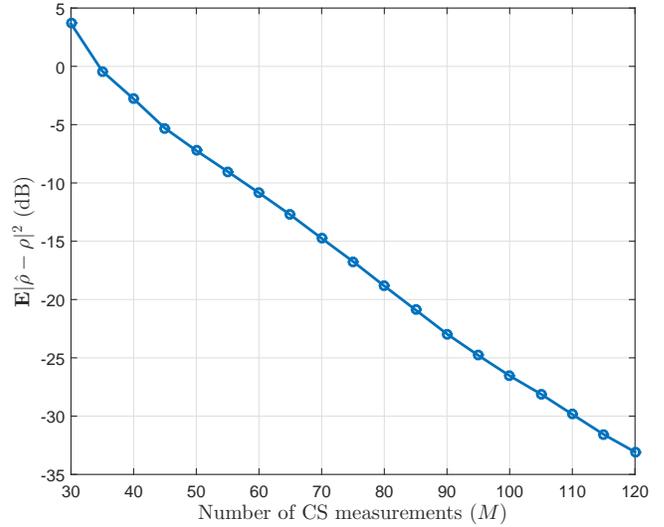


Figure 9: FD-block-CS: Mean Square Error versus M , $J = 3$ and $N = 300$ users.

APPENDIX B

PROOF OF LEMMA 3

Let $\mathbf{H} = \mathbf{B}_T^t \boldsymbol{\Sigma}_z^{-1} \mathbf{B}_T$. 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 in Lemma 4, we can establish that the off-diagonal entries of \mathbf{H} given by

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

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

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

with $d_{j,n} = \mathbf{E} \left\{ (\mathbf{a}_{c,n}^j)^t \boldsymbol{\Sigma}_z^{-1} \mathbf{a}_{c,n}^j \right\}$ reading as

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

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

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

We therefore obtain: $\mathbf{H}^{-1} - \frac{M \mathbf{I}_{JS}}{\text{tr}(\boldsymbol{\Sigma}_z^{-1})} \xrightarrow{M \rightarrow +\infty, a.s.} 0$.

REFERENCES

- [1] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A Survey on 3GPP Heterogeneous Networks," *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, June 2011.
- [2] D. Soldani and S. Dixit, "Wireless Relays for Broadband Access," *IEEE Commun. Mag.*, vol. 46, no. 3, pp. 58–66, Mar 2008.
- [3] J. Laneman, D. Tse, and G. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec 2004.
- [4] J. Sydir and R. Taori, "An Evolved Cellular System Architecture Incorporating Relay Stations," *IEEE Commun. Mag.*, vol. 47, no. 6, pp. 115–121, June 2009.

- [5] T. Cover and A. Gamal, "Capacity Theorems for The Relay Channel," *IEEE Transactions on Information Theory*, vol. 25, no. 5, pp. 572–584, Sep 1979.
- [6] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3037–3063, Sep 2005.
- [7] D. Chen and J. Laneman, "The Diversity-multiplexing Tradeoff for The Multiaccess Relay Channel," *Proc. of the 40th Annual Conf. on Inf. Sciences and Systems*, vol. 47, no. 6, pp. 1324–1328, Mar 2006.
- [8] A. Host-Madsen and J. Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, June 2005.
- [9] M. Hasna and M. S. Alouini, "End-to-end Performance of Transmission Systems with Relays over Rayleigh Fading Channels," *IEEE Trans. on Wireless Communications*, vol. 2, no. 6, pp. 1126–1131, Nov 2003.
- [10] I. Krikidis, H. Suraweera, P. Smith, and C. Yuen, "Full-Duplex Relay Selection for Amplify-and-Forward Cooperative Networks," *IEEE Trans. on Wireless Commun.*, vol. 11, no. 12, pp. 4381–4393, December 2012.
- [11] M. Duarte and A. Sabharwal, "Full-Duplex Wireless Communications Using Off-the-Shelf Radios: Feasibility and First Results," in *Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, Nov 2010, pp. 1558–1562.
- [12] E. Everett, M. Duarte, C. Dick, and A. Sabharwal, "Empowering Full-Duplex Wireless Communication By Exploiting Directional Diversity," in *Conference Record of the Forty Fifth Asilomar Conference on Signals, Systems and Computers*, Nov 2011, pp. 2002–2006.
- [13] T. Riihonen, S. Werner, and R. Wichman, "Mitigation of Loopback Self-Interference in Full-Duplex MIMO Relays," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 5983–5993, Dec 2011.
- [14] R. Tannious and A. Nosratinia, "Spectrally-efficient Relay Selection with Limited Feedback," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1419–1428, Oct 2008.
- [15] A. Tajer and A. Nosratinia, "Opportunistic Cooperation via Relay Selection with Minimal Information Exchange," *Proc. IEEE Int. Symposium on Inf. Theory*, June 2007.
- [16] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A Simple Cooperative Diversity Method Based on Network Path Selection," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 659–672, March 2006.
- [17] D. Michalopoulos, G. Karagiannidis, T. Tsiftsis, and R. Mallik, "An Optimized User Selection Method for Cooperative Diversity Systems," *Proc. IEEE Global Telecommun. Conf.*, vol. 24, no. 3, pp. 1–6, Dec 2006.
- [18] R. Madan, N. Mehta, A. Molisch, and J. Zhang, "Energy-efficient Cooperative Relaying over Fading Channels with Simple Relay Selection," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3013–302, Aug 2005.
- [19] Y. Jing and H. Jafarkhani, "Single and Multiple Relay Selection Schemes and their Achievable Diversity Orders," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1414–1423, March 2009.
- [20] V. Shah, N. Mehta, and R. Yim, "The Relay Selection and Transmission Tradeoff in Cooperative Communication Systems," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 2505–2515, Aug 2010.
- [21] K. Tourki and M.-S. Alouini, "Toward Distributed Relay Selection for Opportunistic Amplify-and-Forward Transmission," *IEEE Veh. Tech. Conf.*, pp. 1–5, May 2011.
- [22] M. Eltayeb, K. Elkhailil, H. Bahrami, and T. Al Naffouri, "Opportunistic Relay Selection with Limited Feedback," *IEEE Transactions on Communications*, vol. PP, no. 99, pp. 1–1, 2015.
- [23] N. Yang, M. Elkhailil, and J. Yuan, "Outage Probability of Multiuser Relay Networks in Nakagami- m Fading Channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2120–2132, Jun 2010.
- [24] X. Liu, X. Zhang, and D. Yang, "Outage Probability Analysis of Multiuser Amplify-and-Forward Relay Network with The Source-to-Destination Links," *IEEE Commun. Lett.*, vol. 15, no. 2, pp. 202–204, Feb 2011.
- [25] Y. Huang, F. Al-Qahtani, C. Zhong, Q. Wu, J. Wang, and H. Alnuweiri, "Performance Analysis of Multiuser Multiple Antenna Relaying Networks with Co-Channel Interference and Feedback Delay," *IEEE Trans. on Commun.*, vol. 62, no. 1, pp. 59–73, Jan 2014.
- [26] N. Yang, M. Elkhailil, and J. Yua, "Impact of Opportunistic Scheduling on Cooperative Dual-Hop Relay Networks," *IEEE Trans. Commun.*, vol. 59, no. 3, pp. 689–694, Mar 2011.
- [27] J. Kim, D. Michalopoulos, and R. Schober, "Diversity Analysis of Multi-User Multi-Relay Networks," *IEEE Trans. Commun.*, vol. 10, no. 7, pp. 2380–2389, July 2011.
- [28] K. Elkhailil, M. Eltayeb, H. Shibli, H. Bahrami, and T. Al-Naffouri, "Opportunistic Relay Selection in Multicast Relay Networks using Compressive Sensing," in *IEEE Global Communications Conference (GLOBECOM)*, Dec 2014, pp. 3126–3131.
- [29] D. Tse, "Optimal Power Allocation Over Parallel Gaussian Broadcast Channels," in *IEEE International Symposium on Information Theory*, Jun 1997.
- [30] C.-J. Chen and L.-C. Wang, "A Unified Capacity Analysis For Wireless Systems With Joint Antenna and Multiuser Diversity in Nakagami Fading Channels," *IEEE Trans. on Commun.*, vol. 54, no. 3, pp. 469–478, March 2006.
- [31] D. Piazza and L. Milstein, "Analysis of Multiuser Diversity in Time-Varying Channels," *IEEE Trans. on Wireless Commun.*, vol. 6, no. 12, pp. 4412–4419, December 2007.
- [32] M. Eltayeb, T. Al-Naffouri, and H. Bahrami, "Compressive Sensing for Feedback Reduction in MIMO Broadcast Channels," *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3209–3222, Sept 2014.
- [33] K. Elkhailil, M. Eltayeb, H. Dahrouj, and T. Al-Naffouri, "Distributed User Selection in Network MIMO Systems with Limited Feedback," to appear in the 2015 IEEE 82nd Vehicular Technology Conference, Sept 2015, Boston, USA.
- [34] M. Eltayeb, K. Elkhailil, A. Mas'ud, and T. Al-Naffouri, "Relay Selection with Limited and Noisy Feedback," to appear in the 2015 IEEE 82nd Vehicular Technology Conference, Sept 2015, Boston, USA.
- [35] E. Candes and Y. Plan, "Near-Ideal Model Selection by l_1 Minimization," *Ann. Statist.*, vol. 37, no. 5A, pp. 2145–2177, 2009.
- [36] R. Baraniuk, V. Cevher, M. Duarte, and C. Hegde, "Model-Based Compressive Sensing," *IEEE Trans. on Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, April 2010.
- [37] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Foundations and Trends in Communications and Information Theory, 2004.
- [38] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge University Press, 2011.
- [39] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [40] A. Hazmi, J. Rinne, and M. Renfors, "Cancellation of Loop Interference With Exponential Profile Using Autocorrelation Method in OFDM Based Systems," in *The Ninth International Conference on Communications Systems*, Sept 2004, pp. 140–144.
- [41] 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 *IEEE Global Telecommunications Conference*, vol. 1, 2000, pp. 167–171 vol.1.
- [42] V. Aggarwal, M. Duarte, A. Sabharwal, and N. Shankaranarayanan, "Full- or Half-Duplex? A Capacity Analysis with Bounded Radio Resources," in *IEEE Information Theory Workshop (ITW)*, Sept 2012, pp. 207–211.
- [43] S. Goyal, P. Liu, S. Panwar, R. Difazio, R. Yang, and E. Bala, "Full Duplex Cellular Systems: Will Doubling Interference Prevent Doubling Capacity?" *IEEE Commun. Magazine*, vol. 53, no. 5, pp. 121–127, May 2015.
- [44] A. Fletcher, S. Rangan, and V. Goyal, "Necessary and Sufficient Conditions for Sparsity Pattern Recovery," *IEEE Trans. on Inf. Theory*, vol. 55, no. 12, pp. 5758–5772, Dec 2009.
- [45] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the Reconstruction of Block-Sparse Signals With an Optimal Number of Measurements," *IEEE Trans. on Sig. Processing*, vol. 57, no. 8, pp. 3075–3085, Aug 2009.
- [46] 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, Aug 2006.
- [47] M. Masood and T. Y. Al-Naffouri, "Support agnostic bayesian matching pursuit for block sparse signals," in *IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada*, 2013, pp. 4643–4647.
- [48] A. Sayed, *Fundamentals of Adaptive Filtering*. John Wiley & Sons, 2003.
- [49] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*. 7th ed., Academic Press, 2007.
- [50] C. Patel, L. Gropop, V. Chande, V. Khaitan, M. Yavuz, and S. Nanda, "Femtocell and Beacon Transmit Power Self-Calibration," *Qualcomm Incorporated*, 2010.
- [51] B. Lindmark, P. Zetterberg, and L. Garcia-Garcia, "Macro-Cellular MIMO with Dual Polarized Base Station Antennas and Different Spacing," in *The Second European Conference on Antennas and Propagation, EuCAP*, Nov 2007, pp. 1–5.