

An Efficient Clustering Algorithm for Device-to-Device Assisted Virtual MIMO

S. Hossein Seyedmehdi and Gary Boudreau

Abstract—In this paper, the utilization of mobile devices (MDs) as decode-and-forward relays in a device-to-device assisted virtual MIMO (VMIMO) system is studied. Single antenna MDs are randomly distributed on a 2D plane according to a Poisson point process, and only a subset of them are sources leaving other idle MDs available to assist them (relays). Our goal is to develop an efficient algorithm to cluster each source with a subset of available relays to form a VMIMO system under a limited feedback assumption. We first show that the NP-hard optimization problem of precoding in our scenario can be approximately solved by semidefinite relaxation. We investigate a special case with a single source and analytically derive an upper bound on the average spectral efficiency of the VMIMO system. Then, we propose an optimal greedy algorithm that achieves this bound. We further exploit these results to obtain a polynomial time clustering algorithm for the general case with multiple sources. Finally, numerical simulations are performed to compare the performance of our algorithm with that of an exhaustive clustering algorithm, and it is shown that these numerical results corroborate the efficiency of our algorithm.

Index Terms—Cooperative diversity, clustering algorithms, virtual MIMO (VMIMO), semidefinite relaxation (SDR).

I. INTRODUCTION

THE use of multiple transmit and receive antennas (MIMO) has been perceived as a promising technique to enhance the spectral efficiency of wireless systems. In practice, due to the size limitation, only one transmit antenna can usually fit inside each mobile device (MD) especially in low cost legacy devices. However, multiple single-antenna MDs can be clustered together to create a virtual MIMO (VMIMO) system. In most current wireless standards, including the long term evolution (LTE) and IEEE 802.11, MDs that are scheduled to transmit comprise only a small fraction of all present MDs, thus leaving numerous idle MDs available to form device-to-device assisted VMIMO systems. One of the new challenges that arises in this realm is that given a large number of MDs, a proper subset of them must be selected to form the VMIMO system (also known as the *clustering* or *user grouping* problem).

Utilizing a cooperating device has been shown to improve the spectral efficiency and diversity while alleviating the outage behaviour [1]–[3]. This cooperation is performed by a second wireless device (the relay) relaying the message of a first wireless device (the source).

There is a rich body of literature on how to select the proper relay(s) to cooperate with the source(s) that can be categorized into three trends: In the earlier works, e.g., [4]–[7], selection of a *single* best relay (for a single source) based on different performance metrics including the best instantaneous channel condition [5] or the trade off between the amplify-and-forward (AF) error probability and the power consumption [7] is studied. Later works have generalized the single relay selection concept to *multiple* relay selection in an effort to find optimal relay selection methods under various assumptions [8]–[16]. Most of these works consider AF relaying (e.g., [8]–[12]). The DF two-hop relaying technique is considered in [14]–[16]. Both [15] and [16] investigate algorithms to select m best relays out of N uniformly distributed available relays for a single source, and obtain approximations on the cumulative distribution function (CDF) of the source spectral efficiency when m relays are selected. In recent works, relay selection for multiple sources is considered in [17]–[21]. AF relaying is considered in [17] where L relays are selected. DF selection is considered in [18]–[21]. In [18] and [19], orthogonal channels are assumed for each source, thus eliminating the effect of the leakage interference. In [18] and [20] a fixed number of relays (one) is selected for each source. The relay selection problem is approached from a *pricing based game model* in [21], [22]. In [21] each relay demands a price (virtual currency) for cooperation, and each source bids a price to recruit relays, and in addition, a *competitive price adjustment process* is discussed. Although selection is not a new concept, a scarcity of studies on clustering (grouping) algorithms from a MIMO perspective is apparent. For instance, most previous works including [17]–[21] consider only single antenna at the receiver and fail to examine the effect of the precoding.

Recently, the problem of clustering algorithms for VMIMO has gained increasing attention since VMIMO is recognized as a promising future trend of communication systems [23]–[25]. The problem of joint grouping and precoding for fixed size VMIMO where precoding weights are continuous is studied in [26]. The uplink pair selection problem (e.g., [27]) is extended to multi user uplink grouping for a single destination in [28]. In this work a fixed number of users are selected (and grouped together) such that a proportional fairness utility is maximized. Nonetheless, the cooperation between nodes is not considered, and each node transmits its individual message.

In this paper, we treat the problem of clustering algorithms for multiple sources by utilizing idle MDs as assisting devices (relays) in a VMIMO setup with limited feedback. In the scenario under consideration, all MDs share the same channel, and they are spatially distributed according to a Poisson

S. H. Seyedmehdi is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, M5S 3G4, Canada (email: hossein@comm.utoronto.ca). G. Boudreau is with Ericsson, Ottawa, ON, Canada (email: gary.boudreau@ericsson.com).

This work was supported in part by Ericsson Canada and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

point process. A source either transmits its message without forming a VMIMO system, or it can be clustered with other idle MDs in a VMIMO setup if the latter act improves its spectral efficiency. When a source participates in the VMIMO configuration, its message is conveyed to its serving access point (AP) in a two-phase DF relaying manner, and unlike some prior studies (e.g., [15], [16]), the direct link is also considered. Moreover, an approximate precoding is performed by which the transmit signal of each MD in the second phase is multiplied by a discrete unity complex number. Due to the limited feedback assumption, unlike [26], these precoding weights are chosen from a finite cardinality codebook.

To the best of our knowledge, this is the first work that studies the network performance as a function of the precoding and density of randomly distributed MDs and proposes an *efficient* clustering algorithm for the VMIMO setup (as modelled in Section II). Towards this goal, in Section III, we approximately solve the NP-hard problem of precoding optimization by a combination of the search space reduction and semidefinite relaxation (SDR). In Section IV, we study a special case with a single source and analytically derive an upper bound on the achievable spectral efficiency. This upper bound encompasses the stochastic geometry of the problem as well as the randomness in channel gains due to the log-normal shadowing. Then we propose a greedy algorithm with quadratic complexity in the number of MDs that achieves this bound, and hence, it is optimal in this case. We then leverage this knowledge in Section V to develop a clustering algorithm for multiple sources. Our proposed algorithm is efficient in the following senses: firstly, it is run in polynomial time; secondly, it eliminates the need for the backhaul communication between APs since it is not required to control the leakage interference; thirdly, it can achieve significant performance gains (cf. Fig. 3, Fig. 5, and Table II). In Section VI, numerical results are provided demonstrating, firstly, that the performance of our algorithm is very close to that of an exhaustive clustering, and secondly, that our algorithm can improve the tradeoff between the spectral efficiency and energy efficiency of the implementation.

II. SYSTEM MODEL AND PROBLEM STATEMENT

A. Notations and Definitions

The terms *rate* and *spectral efficiency* are used interchangeably to indicate the average number of bits per second per hertz (bps/Hz) that can be conveyed through the communication channel. The channel is assumed to be a complex AWGN channel with noise distribution $\mathcal{CN}(0, \sigma_N^2)$. By definition, let $C(x) = \log_2(1 + x)$. The *ceiling* of x is shown by $\lceil x \rceil$. Calligraphic letters (e.g., \mathcal{A}) are used to represent sets. The backslash notation is used to represent the set subtraction, i.e., $\mathcal{A} \setminus \mathcal{B} \triangleq \mathcal{A} - \mathcal{B}$. Vectors are shown by small bold faced letters (e.g., \mathbf{h}), and matrices are shown by capital bold faced letters (e.g., \mathbf{H}). The i -th element of a vector \mathbf{h} is shown by h_i . The conjugate transpose of a matrix is shown by $(\cdot)^*$. The operator $|\cdot|$ is either the absolute value when its operand is a complex (or real) number or the set cardinality if it operates on sets. \mathbf{I}_N represents the unity matrix with dimension $N \times N$.

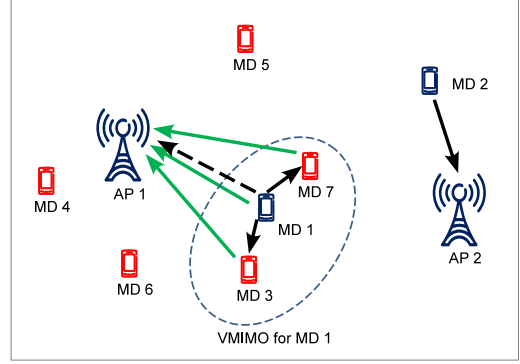


Fig. 1. Illustration of the system model for $M = 2$ destinations (APs) and 7 MDs ($N_{MD} = 7$). In this case, MDs 1 and 2 are scheduled to transmit (sources), and MDs 3 and 7 are assisting the MD 1 in the VMIMO configuration. Black lines represent transmission in the first phase, and green lines represent transmission in the second phase.

B. Distribution of MDs and Scheduling Model

We assume that MDs are randomly distributed in a two dimensional field according to a Poisson point process with rate λ MDs per unit area. The Poisson assumption has been generally accepted as a proper model for the spatial distribution of MDs in wireless networks specially in the presence of a large population of users (cf. [29]–[31]).

Let ξ be a realization of the spatial distribution of MDs, i.e., there are $N_{MD}(\xi)$ MDs on the field given this realization. We further assume that $N_{MD}(\xi)$ remains stationary for a relatively long time. Therefore, we omit ξ hereafter and let MDs be indexed as $1, 2, \dots, N_{MD}$. Without loss of generality, we assume there are M APs, and the scheduler schedules one MD per AP; therefore, there are M sources. Let these M sources be indexed as $1, \dots, M$ where $M \leq N_{MD}$.

C. Structure of the VMIMO and Corresponding Rates

It is assumed that each MD is equipped with one transmit antenna, and each AP is equipped with N_{rx} receive antennas. Each source s either directly transmits to its respective destination without the assistance of other relays or adopts a two-phase transmission where it is assisted by the set of idle MDs (relays) \mathcal{A}_s . In the latter case, in phase one, the source broadcasts its message (codeword), and all MDs k , $k \in \mathcal{A}_s$, decode the message of the source while the AP postpones the decoding. In phase two, the source s and all the MDs k , $k \in \mathcal{A}_s$, transmit the same message as that of the phase one with the precoding, and then, the AP decodes this message by augmenting received signals in phases one and two. In more precise words, if the source s is assisted by other MDs, it transmits the codeword $\mathbf{x}_j^{(1)}$ in phase one and repeats the same codeword in phase two. On the other hand, if the source s is not assisted by other MDs, it transmits $\mathbf{x}_j^{(q)}$ in phase q where $q = 1, 2$. We further assume that the symbols in the a codeword are power normalized, i.e., $E[|x_j|^2] = 1$.

Let $\mathbf{y}_d^{(1)}$ be the received signal in phase one at a specific time at the AP d . We have $\mathbf{y}_d^{(1)} = \sum_{j=1}^M \mathbf{h}_{jd} \sqrt{P_j} x_j^{(1)} + \mathbf{n}_d^{(1)}$ where $\mathbf{y}_d^{(1)}$ is an $N_{rx} \times 1$ vector, \mathbf{h}_{jd} is an $N_{rx} \times 1$ vector

whose elements are the channel gains between the MD j and the AP d , $x_j^{(1)}$ is the power normalized symbol transmitted by the source j in phase one, $\mathbf{n}_d^{(q)}$ is the AWGN vector in phase q , and P_j is the transmit power of the MD j . Also, let $\mathcal{L} = \{1, \dots, L\}$ be the subset of sources that are assisted in the described VMIMO setup. In the second phase, MDs that participate in VMIMO, precode their transmit signal with a unity complex weight factor. Therefore, the received signal at the AP d in phase two can be written as $\mathbf{y}_d^{(2)} = \sum_{j=1}^L \mathbf{H}_{jd} \mathbf{w}_j x_j^{(1)} + \sum_{j=L+1}^M \mathbf{h}_{jd} x_j^{(2)} + \mathbf{n}_d^{(2)}$ where $\mathbf{H}_{jd} = [\mathbf{h}_{jd} \sqrt{P_j} \quad \mathbf{h}_{k_1 d} \sqrt{P_{k_1}} \quad \dots \quad \mathbf{h}_{k_{|\mathcal{A}_j|} d} \sqrt{P_{k_{|\mathcal{A}_j|}}}]$, $k_i \in \mathcal{A}_j$. The column vector \mathbf{w}_j contains $|\mathcal{A}_j|+1$ precoding weights w_{jk} where $w_{jk} \in \{1, w, \dots, w^{N_w-1}\}$ and w is the N_w -th principal root of unity. The augmented received signal at the AP d after two phases is

$$\mathbf{y}_d = \sum_{j=1}^L \begin{bmatrix} \mathbf{h}_{jd} \sqrt{P_j} \\ \mathbf{H}_{jd} \mathbf{w}_j \end{bmatrix} x_j^{(1)} + \sum_{j=L+1}^M \begin{bmatrix} \mathbf{h}_{jd} \sqrt{P_j} \\ 0 \end{bmatrix} x_j^{(1)} + \sum_{j=L+1}^M \begin{bmatrix} 0 \\ \mathbf{h}_{jd} \sqrt{P_j} \end{bmatrix} x_j^{(2)} + \begin{bmatrix} \mathbf{n}_d^{(1)} \\ \mathbf{n}_d^{(2)} \end{bmatrix}. \quad (1)$$

Employing a linear MMSE decoder [32], we first obtain the spectral efficiency of sources that are not assisted by relays. For the source s , $s \in \{L+1, \dots, M\}$, the capacity in phase $q = 1, 2$ can be computed as $c_s^{(q)} = C(P_s \mathbf{h}_{sd}^* \mathbf{K}_{\mathbf{z}^{(q)}}^{-1} \mathbf{h}_{sd})$ where

$$\mathbf{K}_{\mathbf{z}^{(1)}} = \sum_{j=1, j \neq s}^M P_j \mathbf{h}_{jd} \mathbf{h}_{jd}^* + \sigma_N^2 I_{N_{rx}},$$

$$\mathbf{K}_{\mathbf{z}^{(2)}} = \sum_{j=1}^L \mathbf{H}_{jd} \mathbf{w}_j \mathbf{w}_j^* \mathbf{H}_{jd}^* + \sum_{j=L+1, j \neq s}^M P_j \mathbf{h}_{jd} \mathbf{h}_{jd}^* + \sigma_N^2 I_{N_{rx}}.$$

The overall rate of the source s is the average of its rates over two phases, i.e.,

$$r_s = \frac{1}{2} (c_s^{(1)} + c_s^{(2)}), \quad s \in \{L+1, \dots, M\}. \quad (2)$$

The aggregate capacity between the source s when assisted by the relays \mathcal{A}_s in two phases and the destination d , can be written as

$$c_s = C(\tilde{\mathbf{h}}_{sd}^* \mathbf{K}_{\mathbf{z}_s}^{-1} \tilde{\mathbf{h}}_{sd}) \quad (3)$$

where $\tilde{\mathbf{h}}_{sd}^* = [\mathbf{h}_{sd}^* \sqrt{P_s} \quad \mathbf{w}_s^* \mathbf{H}_{sd}^*]$, and

$$\mathbf{K}_{\mathbf{z}_s} = \sum_{j=1, j \neq s}^L \tilde{\mathbf{h}}_{jd} \tilde{\mathbf{h}}_{jd}^* + \sum_{j=L+1}^M \begin{bmatrix} P_j \mathbf{h}_{jd} \mathbf{h}_{jd}^* & 0 \\ 0 & P_j \mathbf{h}_{jd} \mathbf{h}_{jd}^* \end{bmatrix} + \sigma_N^2 I_{2N_{rx}} \quad (4)$$

Therefore, the overall spectral efficiency of the relay assisted source s can be written as

$$r_s = \frac{1}{2} \min \{ \{c_s\} \cup \{r_{sl} : l \in \mathcal{A}_s\} \}, \quad s \in \mathcal{L} \quad (5)$$

where r_{sl} is the rate between the source s and MD l (computed similar to $c_s^{(1)}$), and c_s is the rate in (3). The half factor is present since the assisted source s transmits a repeated version

of its message in two consecutive time slots.

Fig. 1 illustrates the above described system model. In this figure, $\mathcal{L} = \{1\}$, and $\mathcal{A}_1 = \{3, 7\}$.

D. Statement of the Problem

We seek to maximize the harmonic mean utility subject to disjoint set of assisting relays for each source and quantized unity precoding weights. This goal can be mathematically expressed as

$$\max \quad \frac{M}{\sum_{i=1}^M r_i^{-1}} \quad (6a)$$

$$\text{s.t.} \quad w_{jk} \in \{1, w, \dots, w^{N_w-1}\}, \quad j \in \mathcal{L}; \quad (6b)$$

$$\mathcal{A}_i \cap \mathcal{A}_j = \emptyset, \quad i \neq j; \quad (6c)$$

$$\cup_{j=1}^M \mathcal{A}_j \subseteq \{M+1, \dots, N_{MD}\}. \quad (6d)$$

III. AN APPROXIMATION METHOD FOR THE PRECODING PROBLEM

To approach the NP-hard problem of precoding, we isolate the precoding search space for each VMIMO cluster (associated with one source) and show in the following that SDR can be applied to obtain an approximate solution.

The SINR term in (3) can be written as

$$\tilde{\mathbf{h}}_{sd}^* \mathbf{K}_{\mathbf{z}_s}^{-1} \tilde{\mathbf{h}}_{sd} = [\mathbf{h}_{sd}^* \sqrt{P_s} \quad \mathbf{w}_s^* \mathbf{H}_{sd}^*] \mathbf{K}_{\mathbf{z}_s}^{-1} \begin{bmatrix} \mathbf{h}_{sd} \sqrt{P_s} \\ \mathbf{H}_{sd} \mathbf{w}_s \end{bmatrix} = \tilde{\mathbf{w}}_s^* \mathbf{Q}_s \tilde{\mathbf{w}}_s \quad (7)$$

where $\tilde{\mathbf{w}}_s = [1 \quad \mathbf{w}_s^*]^*$, and

$$\mathbf{Q}_s = \begin{bmatrix} \mathbf{h}_{sd}^* \sqrt{P_s} & 0 \\ 0 & \mathbf{H}_{sd}^* \end{bmatrix} \mathbf{K}_{\mathbf{z}_s}^{-1} \begin{bmatrix} \mathbf{h}_{sd} \sqrt{P_s} & 0 \\ 0 & \mathbf{H}_{sd} \end{bmatrix}. \quad (8)$$

The maximization of the SINR in (7) is tantamount to the following maximization problem

$$\max \quad \tilde{\mathbf{w}}_s^* \mathbf{Q}_s \tilde{\mathbf{w}}_s \quad (9a)$$

$$\text{s.t.} \quad \tilde{w}_{si} \in \{1, w, \dots, w^{N_w-1}\}. \quad (9b)$$

The optimization problem expressed in (9) is called the *discrete complex quadratic optimization problem* which belongs to the class of NP-hard problems [33]. However, it can be approximately solved by using the SDR method. In this paper, we first use the CVX package [34] to solve the relaxed version of (9). Next, we adopt a rank-one approximation [35] in addition to the uniform quantization to find the precoding weights from the solution of the semidefinite relaxed problem.

Assuming a limited feedback, each MD participating in the VMIMO setup receives the index of the precoding weight in the precoding codebook, $\{1, w, \dots, w^{N_w-1}\}$, through some feedback mechanism. Since the size of the precoding codebook is N_w , this feedback requires $\log_2(N_w)$ bits.

IV. CLUSTERING FOR A SINGLE SOURCE

In this section we develop an optimal clustering algorithm when there is only one source ($M = 1$) and one destination. This algorithm will be later used in Section V for the general case with multiple sources.

To make computations tractable, we temporarily assume that $N_{rx} = 1$; therefore, the channel gain between users l and m can be expressed as $h_{lm} = |h_{lm}|e^{j\theta_{lm}}$. We further assume that channel gains are influenced by the path loss (PL) and log-normal shadowing¹; therefore,

$$|h_{lm}|^2 = Gd_{lm}^{-\alpha}10^{\sigma_{dB}V_{lm}/10} \quad (10)$$

where G is a constant depending on the operating frequency and the antenna gains, d_{lm} is the distance between users l and m , α is the path loss exponent, σ_{dB} is the shadowing dB-spread, and V_{lm} is a standard Gaussian random variable. In addition, it is assumed that MDs are power controlled and the received SNR at the AP d is constant, i.e.,

$$|h_{ld}|^2 \frac{P_l}{\sigma_N^2} = \gamma, \quad l = 1, \dots, N_{MD}. \quad (11)$$

A. An Upper Bound on the Average Spectral Efficiency

In order to evaluate any algorithm, a performance bound on the achievable spectral efficiency is needed. This performance bound is given in the following theorem.

Theorem 1: For the system described in this section, there exists an upper bound on the average spectral efficiency (bps/Hz) of the source s . For $\sigma_{dB} = 0$, this upper bound is given as

$$C(\gamma) + \int_{C(\gamma)} \sum_{k \geq k_r} \frac{[\lambda \Psi_s(r)]^k}{k!} e^{-\lambda \Psi_s(r)} dr \quad (12)$$

where λ is the average MD density (number of MDs per unit area), and $\Psi_s(r)$ is the r -achievability area and given by

$$\Psi_s(r) = \pi \left[\frac{GP_s/\sigma_N^2}{2^{2r} - 1} \right]^{2/\alpha}, \quad (13)$$

and k_r is the r -necessary number of relays and given by

$$k_r = \left\lceil \sqrt{\frac{2^{2r} - 1 - \gamma}{\gamma}} - 1 \right\rceil. \quad (14)$$

Proof: See Appendix A. ■

There are a number of important implications associated with Theorem 1. Firstly, the second term ($\int(\cdot)dr$) in (12) represents the average improvement in the spectral efficiency achieved by the formation of the VMIMO as compared with the baseline spectral efficiency $C(\gamma)$ without VMIMO. In other words, if the source can not benefit from the assistance of other MDs, the second term in (12) will be zero. Moreover, when the mean channel gains are influenced by PL, the r -achievability area and r -necessary number of relays can be used to simplify the clustering in a practical implementation when the geographical location of MDs is known. We state these results in the following two corollaries. The proofs follow directly from the proof of Theorem 1, and we skip them.

¹Note that the channel gain is usually composed of three major components: path loss, slow fading (shadowing) modelled as a log-normal random variable, and fast fading modelled as a Rayleigh random variable. However, the mean channel gain is only affected by the path loss and shadowing.

Corollary 1: Let R_s be the maximum average spectral efficiency expressed in (12), and the channel gains are affected by the PL. The number of relays necessary to achieve R_s is given by the R_s -necessary number of relays, k_{R_s} .

Corollary 2: Let R_s be the maximum average spectral efficiency expressed in (12), and the channel gains are affected by the PL. The expected area in which assisting MDs are located is bounded by a disk centred at MD s with radius $\sqrt{\Psi_s(R_s)/\pi}$.

For the general case with the log-normal shadowing, i.e., $\sigma_{dB} > 0$, an asymptotic upper bound can be obtained. The derivation of this bound is rather involved; we leave the details of this derivation to Appendix B and state the results in the following lemma.

Lemma 1: For the system described in this section and some positive δ , assume that relays are located within a disk with radius d_{\max} ($d_{\max} > \delta > 0$) centred at the source s . Let the probabilities $\pi_k(r, \delta)$ be calculated according to (30). When $\delta \rightarrow 0$, the average spectral efficiency of the source s can be upper bounded as

$$C(\gamma) + \int_{C(\gamma)} \left[\sum_{k \geq k_r} \pi_k(r, \delta) \right] dr. \quad (15)$$

Note that the mean of the shadowing term in (10) is greater than one, i.e., $E[10^{\sigma_{dB}V_{lm}/10}] \geq 1$. In other words, it is expected that the log-normal shadowing increases the connectivity of the network, consistent with the results shown in [36]. This fact results in the bound for the average spectral efficiency in Lemma 1 being greater than that in Theorem 1. This difference is illustrated in Fig. 2, and as can be seen, the log-normal shadowing increases the average spectral efficiency. It is also illustrated that these bounds are tight for infinite number of precoding weights.

B. An Optimal $\mathcal{O}(N_{MD}^2)$ Algorithm

For a fixed realization of MDs, let s be the source, and let the set of available idle MDs (relays) be $\mathcal{I} = \{1, \dots, N_{MD}\} \setminus \{s\}$. Denote as Algorithm 1, an algorithm that can perform optimal clustering when $N_w = \infty$. Under this condition, an optimal precoding can be performed by multiplying the signal x_l by the precoding weight $e^{-j\theta_{ld}}$, thus cancelling the phase rotation caused by the channel.

There are four major steps in Algorithm 1: In the first step, the candidate relays are discovered. Candidate relays are the ones whose rate to the source is greater than twice as much as that of the source to the AP. In the second step, these candidate relays are sorted in a descending ranking based on their link rate to the source. In the third step, these ranked candidate relays are added to the cluster one by one if with each addition, the spectral efficiency of the source increases. In the last step, it is decided whether the VMIMO system is formed or the source transmits in a one-phase manner.

The precoding step in Algorithm 1 for scalar channel gains can be performed in $\mathcal{O}(N_{MD})$ steps. Therefore, the overall complexity of the algorithm would be $\mathcal{O}(N_{MD}^2)$.

Proposition 1: Algorithm 1 selects the best subset of MDs to cluster with the source, and therefore, on average it can achieve the bound given in Theorem 1 for $N_w = \infty$.

Algorithm 1 Clustering for the system described in Sec. IV

```

 $r_s \leftarrow C(\gamma);$   $\triangleright$  Link rate without VMIMO
 $\mathcal{E}_s \leftarrow \{l : l \in \mathcal{I}, r_{sl} > 2r_s\}$   $\triangleright$  Candidate relay discovery
 $\mathcal{E}_s^{\text{sorted}} \leftarrow \text{Sort } \mathcal{E}_s \text{ descending w.r.t. } r_{sj}, j \in \mathcal{E}_s$ 
 $\mathcal{A}_s \leftarrow \emptyset; r_{\text{old}} \leftarrow 0$ 
for  $j$  in  $\mathcal{E}_s^{\text{sorted}}$  do
   $\mathcal{A}_s \leftarrow \mathcal{A}_s \cup \{j\}$ 
   $r_{\min} \leftarrow \min \{r_{sl} : l \in \mathcal{A}_s\}$ 
   $p(\mathbf{w}) \triangleq \|\mathbf{h}_{sd}\|^2 \frac{P_s}{\sigma_N^2} + \frac{1}{\sigma_N^2} \left\| \sum_{i \in \{s\} \cup \mathcal{A}_s} \mathbf{h}_{id} w_i \sqrt{P_i} \right\|^2$ 
   $\gamma_{\text{AP}} \leftarrow \max_{\mathbf{w}} p(\mathbf{w})$   $\triangleright$  Precoding
   $r_{\text{new}} \leftarrow \frac{1}{2} \min \{C(\gamma_{\text{AP}}), r_{\min}\}$ 
  if  $r_{\text{new}} > r_{\text{old}}$  then
     $r_{\text{old}} \leftarrow r_{\text{new}}$ 
  else
     $\mathcal{A}_s \leftarrow \mathcal{A}_s \setminus \{j\}$ 
  end if
end for
if  $r_{\text{old}} < r_s$  then
   $\mathcal{A}_s \leftarrow \emptyset$   $\triangleright$  No VMIMO
end if

```

Proof: See Appendix C. ■

Remark 1: Algorithm 1 can be considered to belong to the popular category of *greedy* algorithms [37], i.e., it chooses the best relay at each step. In fact, all algorithms (including the ones in [8], [9], [15], [16]) that sweep through a sorted list of items to select the best set of items fall under the category of greedy algorithms. However, there are four major differences between these works and our work in terms of the selection algorithm: firstly, in our work an adaptive number of relays is selected whereas the number of relays is fixed in [16]; secondly, by considering the effect of the direct transmission, our algorithm chooses no relay when the source can not benefit from relays in terms of achievable rates whereas in [9] and [15], the assumption is that source has to select at least one relay; thirdly, unlike [15], [16], Algorithm 1 is not required to sort *all* the relays by establishing a necessary condition for the candidate relays in the discovery step of the algorithm, thus reducing the complexity; lastly, unlike [8] and [9], our greedy algorithm is optimal.

Fig. 2 compares the average spectral efficiency (bps/Hz) vs. λ (MDs/m²) for different received SNRs and different schemes as described in Table I. For this figure, Algorithm 1 is performed over 100 random trials of spatial distribution of MDs. As can be seen, the spectral efficiency gained by Algorithm 1 overlaps with the upper bound on the average spectral efficiency given in Theorem 1 and Lemma 1 when exact phase matching is performed, i.e., $N_w = \infty$ for graphs labeled (ii) and (iv). This observation further corroborates the optimality of Algorithm 1 stated in Proposition 1. The precoding for this simulation (in graphs (v)–(vi)) is performed as described in Section III. It can be speculated that this approximate precoding method improves the performance as compared to the case without the precoding ($N_w = 1$). Furthermore, the improvement in the spectral efficiency due to the VMIMO is more significant for poor performing MDs. For

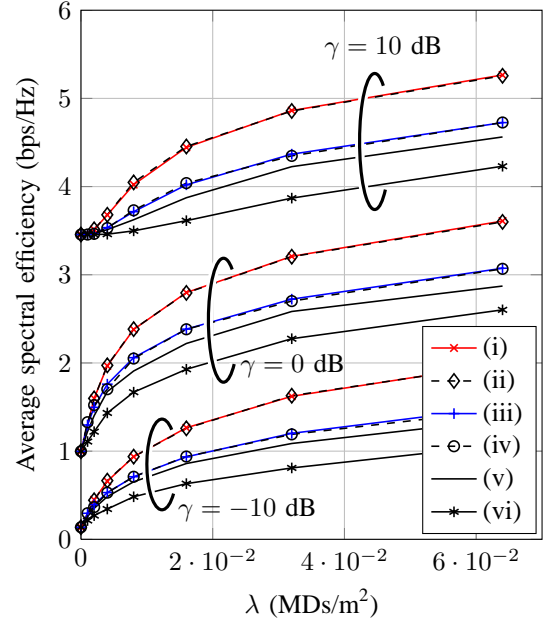


Fig. 2. Average spectral efficiency vs. density of the MDs for various received SNRs. Curves are grouped for each received SNR γ , and the labels in the legend are explained in Table I. Note that $\lambda = 0$ is treated as a special case in which there is only one source and no other MDs. $d_{\max} = 25m$ for (i) and (ii), and $\delta = 0.05$ for (i).

TABLE I
DESCRIPTION OF THE LEGENDS IN FIG. 2

Label	Method	σ_{dB}	N_w
(i)	Lemma 1	8	N/A
(ii)	Algorithm 1	8	∞
(iii)	Theorem 1	0	N/A
(iv)	Algorithm 1	0	∞
(v)	Algorithm 1	0	8
(vi)	Algorithm 1	0	1

instance, for $\lambda = 0.01$ MDs/m², the VMIMO algorithm can boost the spectral efficiency of MDs with $\gamma = -10$ dB more than 400% as compared with the baseline spectral efficiency of $C(\gamma) = 0.14$ bps/Hz whereas this improvement is about 10% for MDs with $\gamma = 10$ dB.

V. CLUSTERING FOR MULTIPLE SOURCES

In this section, we extend the algorithm developed in Section IV to the general case with multiple sources and develop Algorithm 2. Algorithm 1 is used in Algorithm 2 as a sub-algorithm with the following modifications: firstly, $C(\cdot)$ is obtained by the general formula given in (2) and (3); secondly, the precoding is performed according to Section III.

Unlike Algorithm 1, we were unable to prove an optimality bound for Algorithm 2. However, through extensive simulations, we later illustrate that Algorithm 2 exhibits a near optimal performance when compared to that of an exhaustive clustering algorithm as far as the harmonic mean utility metric is concerned.

To motivate why a simple source sorting results in a near optimal performance, it should be noted that one important

Algorithm 2 Clustering algorithm for multiple sources

```

 $\mathcal{M} \leftarrow \{1, \dots, M\}$ 
 $\mathcal{I} \leftarrow \{M+1, \dots, N_{MD}\}$   $\triangleright$  Available idle MDs
 $\mathcal{M}^{\text{sorted}} \leftarrow \text{Sort } \mathcal{M} \text{ ascending w.r.t. } c_s, s \in \mathcal{M}$ 
for  $s$  in  $\mathcal{M}^{\text{sorted}}$  do
    Perform Algorithm 1 for MD  $s$  and idle MD set  $\mathcal{I}$ 
     $\mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{A}_s$ 
end for

```

property of the harmonic mean is that it is limited by the smallest term². This property implies that an increase in the smallest term can potentially lead to a significant increase in the harmonic mean. Therefore, in Algorithm 2, *underprivileged* sources are served first so as to enable them to benefit from the most available resources.

From an algorithmic perspective, we can distinguish Algorithm 2 from previous works on DF relay selection for multiple sources regarding three aspects. Firstly, our algorithm selects an adaptive number of relays for each source while in [18] and [20], one relay is selected for each source. As we argued, enforcing the source to select relays is not always optimal in terms of the spectral efficiency. Secondly, even though Algorithm 2 is oblivious to the leakage interference, the users are assumed to share the same channel. To isolate the relay selection for each user, however, in [18] and [19], it is assumed that sources utilize orthogonal channels. Thirdly, in [19]–[21], conflicts in selection are resolved by deploying some variation of the *message passing* procedure [38] (also known as auction rounds). For instance, [19] and [20] require up to M and $M - 1$ rounds of iteration, respectively. This number can be even larger in [21] depending on system parameters. In other words, the number of iteration rounds scale with the number of sources, and therefore, these algorithms may not be implementable in a large scale scenario where the channels are dynamically and rapidly changing. This problem is addressed in our work by prioritizing sources and requiring only one round of iteration. Consequently, we show that a *simple* and *agnostic* algorithm can perform nearly optimal when the harmonic mean utility metric is considered, and it may not be necessary to resort to a complex and time taking solution.

VI. NUMERICAL RESULTS

A. Simulation Setup and Parameters

In our simulations, we use the 3GPP channel model for the indoor environment [39, Sec. A.2]. According to this model the PL (in dB) for 2 GHz carrier frequency is given by $103.4 + 24.2 \log(d)$ where d is the distance in km. The slow fading is modelled by the log-normal shadowing with dB-spread $\sigma_{dB} = 8$. Antennas are assumed to be omnidirectional with 0 dB antenna gain. The fast fading is assumed to be Rayleigh fading modelled as a Gaussian complex random variable with variance $1/2$ per real dimension. The noise power is $\sigma_N^2 = -101$ dBm for 20 MHz bandwidth. The

maximum transmit power of MDs is limited to $P_{\max} = 20$ dBm. MDs are assumed to be power controlled, and the power control algorithm is assumed to average out the effect of Rayleigh fading, i.e., its decisions are only based on the PL and shadowing. As a result of this power control, each MD's power is adjusted such that the received power at the serving AP is -80 dBm excluding the interference. The number of receive antennas at the AP is $N_{rx} = 4$. The field size is $100\text{m} \times 100\text{m}$. We assume that there are five APs, and therefore, at each instance, there are five sources. Furthermore, to simulate the environments with interference, we assume there is an aggressor network with average density of 10^{-3} sources per square meter with a similar power control algorithm. The clustering algorithm has no control over this aggressor network.

For the sake of a better visualization in our simulation results, we use the effective SINR (in dB) instead of the spectral efficiency. In other words, if r_k is the spectral efficiency of the source k , the corresponding effective SINR, SINR_{eff} , is calculated as

$$\text{SINR}_{\text{eff}} = 10 \log_{10} (2^{r_k} - 1). \quad (16)$$

For each λ or N_w , the reported simulation results are averaged over 1000 trials. These trials encompass the random spatial distribution of MDs and APs, random log-normal shadowing, random Rayleigh fading, and random scheduling.

B. Discussion

The performance of Algorithm 2 is illustrated in Figs. 3, 4, 5, Table II, and Table III. In these illustrations, λ represents the density of the network (MDs/m²), and N_w represents the number of precoding choices (there is no precoding when $N_w = 1$). For low densities of MDs, these performances are compared to that of an exhaustive clustering. However, since the run time of the exhaustive clustering grows exponentially with the density of MDs, for high density networks, it is not computationally feasible to run the exhaustive algorithm.

Fig. 3 shows the improvement in the SINR_{eff} of the sources after the formation of the VMIMO system vs. the baseline SINR_{eff} of sources without the VMIMO. As can be seen, the spectral efficiency of *poor* performing sources is significantly increased whereas the spectral efficiency of the *strong* sources with high SINR_{eff} is slightly degraded. For instance, when there is one MD in every 15 m² on average ($\lambda = 0.064$), sources with $\text{SINR}_{\text{eff}} = -10$ dB can gain more than 7 dB improvement in their SINR_{eff} while sources with $\text{SINR}_{\text{eff}} = 10$ dB suffer less than 1 dB. This compromise, nevertheless, leads to an overall improvement (cf. Table II).

Fig. 4 illustrates the average number of assisting MDs (relays) due to the VMIMO vs. the baseline SINR_{eff} of sources without the VMIMO. As can be seen, the number of assisting MDs tend to increase when the precoding is performed because the aggregate uplink rate in (3) is increased due to this precoding. Hence, considering (5), the number of assisting relays can be increased.

Fig. 5 demonstrates the CDF of the SINR_{eff} for different densities of MDs and precoding resolutions. As can be seen,

² To clarify, consider the following inequality that holds for positive numbers r_1, r_2, \dots, r_M : $M / (\sum_{i=1}^M r_i^{-1}) \leq M \min_{1 \leq i \leq M} r_i$.

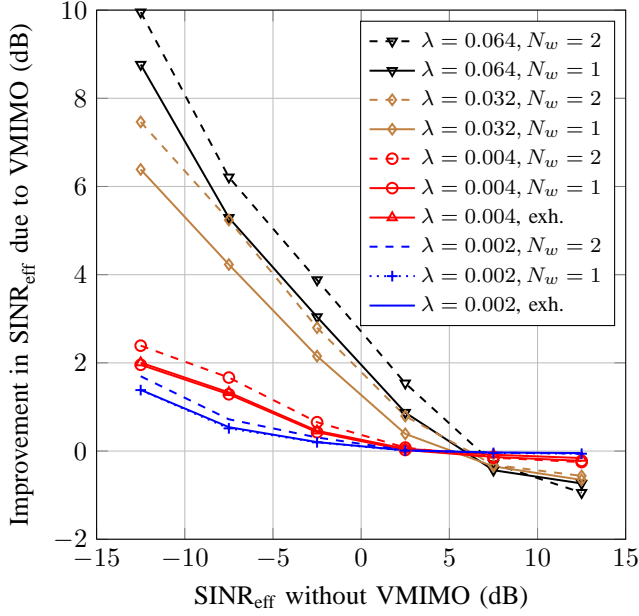


Fig. 3. The improvement of the SINR_{eff} after the formation of VMIMO using Algorithm 2 vs. the baseline SINR_{eff} of sources without VMIMO. λ represents the MD density (MDs/m^2), and N_w represents the number of precoding choices (there is no precoding for $N_w = 1$).

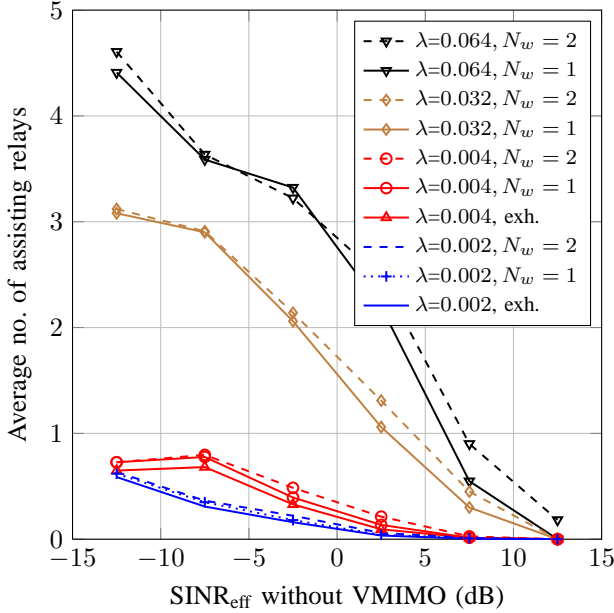


Fig. 4. Average number of assisting MDs (relays) using Algorithm 2 vs. the baseline SINR_{eff} of sources without VMIMO. λ represents the MD density (MDs/m^2), and N_w represents the number of precoding choices (there is no precoding for $N_w = 1$).

the shift of the SINR_{eff} towards higher values on the right, increases as the density of MDs increases. Also, the results clearly illustrate gains of up to 7 dB for the worst 5-10% of users and a gain of 3 dB on average.

The effect of the precoding codebook size (N_w) is illustrated in Tables II and III. In our simulation, before implementing the VMIMO, the overall harmonic mean of the spectral efficiency

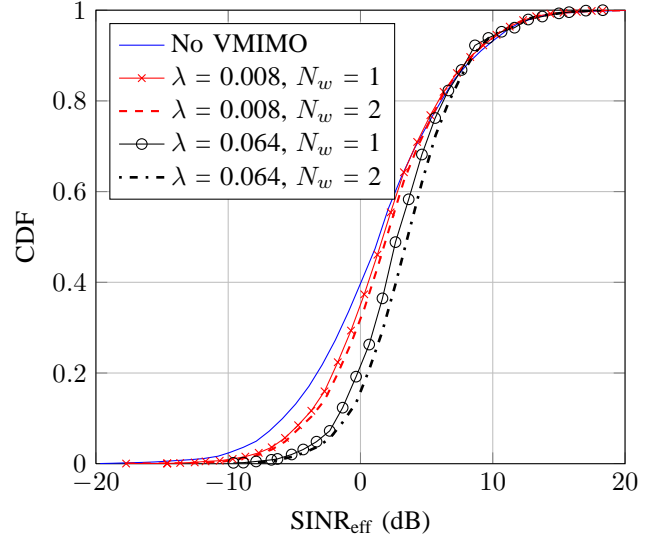


Fig. 5. Cumulative distribution function of the SINR_{eff} for different densities of MDs and precoding resolutions. λ represents the MD density (MDs/m^2), and N_w represents the number of precoding choices (there is no precoding for $N_w = 1$).

is 0.83 (bps/Hz), and the \bar{E}_b is $69.5 \mu\text{J/b}$. As can be seen, one bit of feedback is enough to provide the majority of the gains. Also, juxtaposing Table II and Table III reveals compelling results: firstly, by applying Algorithm 2, the spectral efficiency and energy efficiency are both improved for low densities of MDs; secondly, for high densities of MDs, Algorithm 2 provides a tradeoff, i.e., the harmonic mean of the spectral efficiency is improved at the cost of higher energy per bit. However, the more accurate the precoding is, the higher the gain in the harmonic mean of the spectral efficiency is, and the lower the energy price is.

VII. CONCLUDING REMARKS

In this paper, we developed an efficient clustering algorithm for the device-to-device assisted VMIMO systems with limited feedback, and investigated the effect of the approximate precoding (transmit beamforming) and the MD density on the performance of this VMIMO system. As observed in the numerical simulations, the VMIMO system can significantly boost the performance of *weak* sources while only slightly degrading that of *strong* ones, thus leading to a considerable overall performance increment. These observations would suggest an approach of focusing the user clustering on weak users and not necessarily on strong users. In addition, it was shown that a single bit of feedback for the precoding weight is sufficient to provide the majority of the gain (Tables II and III). Nonetheless, only the repetition-based cooperative diversity gain has been exploited in this work; utilizing space-time-coded cooperative diversity or the multiplexing gain and studying the tradeoff between the diversity and multiplexing gains is a promising area for future investigations.

VIII. ACKNOWLEDGEMENT

The first author would like to thank Mr. Adrien Comeau of Ericsson Canada for his constructive feedbacks on this subject.

TABLE II
IMPROVEMENT IN THE HARMONIC MEAN OF THE SPECTRAL EFFICIENCY
FOR THE COVERAGE THRESHOLD $\text{SINR}_{\text{EFF}} = -10$ dB.

λ	Algorithm 2		
	$N_w = 1$	$N_w = 2$	$N_w = 64$
0.002	2%	4%	4%
0.008	13%	18%	18%
0.064	43%	57%	58%

TABLE III
INCREASE IN THE AVERAGE CONSUMED ENERGY PER BIT \bar{E}_b (J/b) DUE TO
THE VMIMO SETUP.

λ	Algorithm 2		
	$N_w = 1$	$N_w = 2$	$N_w = 64$
0.002	0.2%	-6%	-6%
0.008	11%	3%	3%
0.064	81%	69%	68%

APPENDIX A PROOF OF THEOREM 1

In what follows, denote by $\mathcal{E}_s(r)$ (or \mathcal{E}_s for brevity) the set of all *eligible* relays to participate as a DF relay for the source s when its spectral efficiency is set to r . In other words, \mathcal{E}_s includes relays whose link budget to the source s allow an achievable rate greater than $2r$. The condition for the relay l to be in the set \mathcal{E}_s can be written as

$$l \in \mathcal{E}_s(r) \leftrightarrow C \left(|h_{sl}|^2 \frac{P_s}{\sigma_N^2} \right) \geq 2r. \quad (17)$$

To derive the ASE, the probability of the event $\{R_s > r\}$ for a given r and a realization of the spatial relay distribution and shadowing needs to be obtained where R_s is the overall spectral efficiency of the source s . To achieve this goal two different cases must be considered:

1. $r \leq C(\gamma)$: In this case, the source transmits its message in one phase transmission without recruiting relays. Therefore, in this case, $\Pr\{R_s > r\} = \Pr\{c_{sd} > r\} = 1$.
2. $r > C(\gamma)$: In this case, a non-zero probability of $\{c_{sd} > r\}$ is feasible only if a spectral efficiency increment is achieved by employing relays in the two-phase transmission.

Considering the second case hereafter, by applying the law of total probability, it can be written

$$\Pr\{R_s > r\} = \sum_{k \geq 1} \Pr\{R_s > r | |\mathcal{E}_s| = k\} \Pr\{|\mathcal{E}_s| = k\}.$$

Considering (5), we can write

$$\Pr\{R_s > r | |\mathcal{E}_s(r)| = k\} = \Pr\{c_s > 2r | |\mathcal{E}_s(r)| = k\} \quad (18)$$

where

$$\begin{aligned} c_s &= C \left(|h_{sd}|^2 \frac{P_s}{\sigma_N^2} + \left| \sum_{i \in \{s\} \cup \mathcal{E}_s} h_{id} w_i \sqrt{P_i} \right|^2 / \sigma_N^2 \right) \\ &= C \left(\gamma + \gamma \left| \sum_{i \in \{s\} \cup \mathcal{E}_s} e^{j\hat{\theta}_i} \right|^2 \right). \end{aligned} \quad (19)$$

$\hat{\theta}_i = \theta_i + \theta_{w_i}$ is the phase of the channel gain of MD i to the AP when its phase is shifted by the preceding weight w_i . It can be seen from (19) that c_s is maximized when $\left| \sum_{i \in \{s\} \cup \mathcal{E}_s} e^{j\hat{\theta}_i} \right|$ is maximized. The maximum value of the latter quantity is $k + 1$ when all $\hat{\theta}_i$'s are equal for $i \in \{s\} \cup \mathcal{E}_s$. Given this assumption, we can write

$$\begin{aligned} \Pr\{c_s > 2r | |\mathcal{E}_s(r)| = k\} \\ &\leq \Pr\{C(\gamma[1 + (k+1)^2]) > 2r | |\mathcal{E}_s(r)| = k\} \\ &= \Pr\{k \geq k_r | |\mathcal{E}_s(r)| = k\} \end{aligned} \quad (20)$$

where k_r can be easily computed and is given in the Theorem 1. Moreover, the event $\{|\mathcal{E}_s(r)| = k\}$ can be translated as the event such that k MDs (indexed by l , $l \in \{1, \dots, N_{MD}\} \setminus \{s\}$) are spatially distributed such that

$$C \left(\frac{G}{d_{sl}^\alpha} 10^{\sigma_{dB} V_l / 10} \frac{P_s}{\sigma_N^2} \right) \geq 2r \quad (21)$$

where V_l is a normal Gaussian random variable. The inequality in (21) can be rearranged as

$$d_{sl} e^{-\sigma V_l} \leq d_r \quad (22)$$

where $\sigma = 0.1 \ln(10) \sigma_{dB} / \alpha$, and

$$d_r = \left[\frac{GP_s / \sigma_N^2}{2^{2r} - 1} \right]^{1/\alpha}.$$

Therefore, the eligibility condition in (17) can be simplified as

$$l \in \mathcal{E}_s(r) \leftrightarrow d_{sl} e^{-\sigma V_l} \leq d_r. \quad (23)$$

For $\sigma_{dB} = 0$ (no shadowing), let $\Psi_s(r)$ be the area of the disc centred at the location of the source s with radius d_r such that (21) holds ($\Psi_s(r)$ defined in the Theorem 1). The probability of k Poisson arrivals in the area $\Psi_s(r)$ can be obtained as

$$\Pr\{|\mathcal{E}_s(r)| = k\} = \frac{[\lambda \Psi_s(r)]^k}{k!} e^{-\lambda \Psi_s(r)}. \quad (24)$$

The probability in (20) is either zero when $k < k_r$ or one when $k \geq k_r$. Therefore, the complementary CDF of the rate R_s can be upper bounded as

$$\begin{aligned} \Pr\{R_s > r\} &\leq \sum_{k \geq k_r} \Pr\{|\mathcal{E}_s(r)| = k\} \\ &= \sum_{k \geq k_r} \frac{[\lambda \Psi_s(r)]^k}{k!} e^{-\lambda \Psi_s(r)}. \end{aligned} \quad (25)$$

Since R_s is a non-negative random variable, we can express the expected value of R_s as

$$\begin{aligned} E[R_s] &= \int_0^\infty \Pr\{R_s > r\} dr \\ &= \int_0^{C(\gamma)} dr + \int_{C(\gamma)}^\infty \Pr\{R_s > r\} dr \\ &\leq C(\gamma) + \int_{C(\gamma)}^\infty \sum_{k \geq k_r} \frac{[\lambda \Psi_s(r)]^k}{k!} e^{-\lambda \Psi_s(r)} dr. \end{aligned} \quad (26)$$

APPENDIX B PROOF OF LEMMA 1

For the case with the log-normal shadowing ($\sigma_{dB} > 0$), we adopt an asymptotic method to find $\Pr\{|\mathcal{E}_s| = k\}$ in (24).

Assume that relays are randomly located according to a shrinking Bernoulli process on a disk with radius d_{\max} centred at the source. For this process, $2\pi\lambda\delta d_i$ is the probability that there exists one user in the distance interval of $[d_i - \delta, d_i]$ of the source node where $d_i = i\delta$. This process asymptotically approaches a Poisson process with rate λ as $\delta \rightarrow 0$ [40]. Let $n_{\max} = \lfloor d_{\max}/\delta \rfloor$. For $i = 1, \dots, n_{\max}$, let

$$E_i \triangleq \{\text{exists one relay } l \text{ s.t. } d_{sl}e^{-\sigma V_l} \in [(i-1)\delta, i\delta]\},$$

$$E'_i \triangleq \{\text{exists one relay at } d_i\}.$$

Furthermore, by definition, let

$$p_{j \rightarrow i} = \Pr\{j\delta e^{-\sigma V} \in [(i-1)\delta, i\delta] | E'_j\} \quad (27)$$

for $i, j \in \{1, 2, \dots, n_{\max}\}$ where V is a normal Gaussian random variable. This probability can be computed as

$$p_{j \rightarrow i} \stackrel{(a)}{=} \Pr\{(i-1)\delta \leq j\delta e^{\sigma V} < i\delta\}$$

$$= \frac{1}{2}\text{erf}\left(\frac{1}{\sqrt{2}\sigma} \ln \frac{i}{j}\right) - \frac{1}{2}\text{erf}\left(\frac{1}{\sqrt{2}\sigma} \ln \frac{i-1}{j}\right) \quad (28)$$

where (a) holds since V and $-V$ have the same distributions. The probability of the event E_i, p_i , can be written as

$$p_i = \sum_{j=1}^{n_{\max}} \Pr\{j\delta e^{-\sigma V} \in [(i-1)\delta, i\delta] | E'_j\} \Pr\{E'_j\}$$

$$= 2\pi\lambda\delta^2 \sum_{j=1}^{\lfloor d_{\max}/\delta \rfloor} j p_{j \rightarrow i}. \quad (29)$$

Now, the event $\{|\mathcal{E}_s(r)| = k\}$ is tantamount to the event of k successes out of n YES/NO trials with probability of successes p_1, p_2, \dots, p_n which has a Poisson binomial distribution ($n = \lfloor d_r/\delta \rfloor$). Let $\Pr\{|\mathcal{E}_s(r)| = k\} = \lim_{\delta \rightarrow 0} \pi_k(r, \delta)$. Using a recursive method [41], we have

$$\pi_k(r, \delta) = \begin{cases} \prod_{i=1}^n (1 - p_i), & k = 0, \\ \frac{1}{k} \sum_{j=1}^k (-1)^{j-1} \pi_{k-j}(r, \delta) T_j, & k > 0 \end{cases} \quad (30)$$

where

$$T_j = \sum_{i=1}^n \left(\frac{p_i}{1 - p_i} \right)^j.$$

APPENDIX C PROOF OF PROPOSITION 1

We first claim that in Algorithm 1, r_{\min} is a non-increasing function of $|\mathcal{A}_s|$. This claim can be verified as follows: considering that the set of available MDs for assistance, $\mathcal{E}_s^{\text{sorted}}$, is sorted with respect to the link rates R_{sl} , $l \in \mathcal{E}_s^{\text{sorted}}$, each addition to the \mathcal{A}_s either limits r_{\min} to a lower value or leaves it intact.

Moreover, we claim that in Algorithm 1, γ_{AP} is a non-decreasing function of $|\mathcal{A}_s|$ if the precoding resolution is high enough ($N_w = \infty$), and this claim can be easily verified by observing that in every step of the **for** loop, γ_{AP} either remains the same or becomes larger.

Considering these two claims, since r_{\min} is increasing through the loop and γ_{AP} is decreasing through the loop, there is an optimal point in the loop that adding or deleting any MD to or from \mathcal{A}_s would only decrease r_{new} .

REFERENCES

- [1] J. Laneman and G. Wornell, "Distributed space-time-coded protocols for exploiting cooperative diversity in wireless networks," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2415–2425, Oct. 2003.
- [2] 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.
- [3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity. part I. system description," *IEEE Trans. Commun.*, vol. 51, no. 11, pp. 1927–1938, Nov. 2003.
- [4] A. Ibrahim, A. Sadek, W. Su, and K. Liu, "Cooperative communications with relay-selection: when to cooperate and whom to cooperate with?" *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, Jul. 2008.
- [5] A. Bletsas, A. Khisti, D. Reed, and A. Lippman, "A simple cooperative diversity method based on network path selection," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 659–672, Mar. 2006.
- [6] A. Adinoyi, Y. Fan, H. Yanikomeroglu, H. Poor, and F. Al-Shaalan, "Performance of selection relaying and cooperative diversity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5790–5795, Dec. 2009.
- [7] Y. Zhao, R. Adve, and T. J. Lim, "Improving amplify-and-forward relay networks: Optimal power allocation versus selection," in *IEEE Intl. Symp. Inf. Theory (ISIT)*, Jul. 2006, pp. 1234–1238.
- [8] D. Michalopoulos, G. Karagiannidis, T. Tsiftsis, and R. Mallik, "An optimized user selection method for cooperative diversity systems," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, 2006, pp. 1–6.
- [9] 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.
- [10] S. Ikki and M. Ahmed, "Performance analysis of generalized selection combining for amplify-and-forward cooperative-diversity networks," in *IEEE Intl. Conf. Commun. (ICC)*, Jun. 2009, pp. 1–6.
- [11] K. S. Hwang, Y. C. Ko, and M. S. Alouini, "Performance analysis of incremental opportunistic relaying over identically and non-identically distributed cooperative paths," *IEEE Trans. Wireless Commun.*, vol. 8, no. 4, pp. 1953–1961, Apr. 2009.
- [12] W. Guo, J. Liu, L. Zheng, Y. Liu, and G. Zhang, "Performance analysis of a selection cooperation scheme in multi-source multi-relay networks," in *Intl. Conf. Wireless Commun. Signal Process. (WCSP)*, 2010, pp. 1–6.
- [13] V. Shah, N. Mehta, and R. Yim, "Analysis, insights and generalization of a fast decentralized relay selection mechanism," in *IEEE Intl. Conf. Commun. (ICC)*, 2009, pp. 1–6.
- [14] S. Vakil and B. Liang, "Cooperative diversity in interference limited wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 8, pp. 3185–3195, Aug. 2008.
- [15] S. Nam, M. Vu, and V. Tarokh, "Relay selection methods for wireless cooperative communications," in *Conf. Inf. Sci. Systems (CISS)*, Mar. 2008, pp. 859–864.
- [16] C. H. Yu and O. Tirkkonen, "Opportunistic multiple relay selection with diverse mean channel gains," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 885–891, Mar. 2012.
- [17] J. Wu, Y. Zhang, M. Amin, and M. Uysal, "Multiple-relay selection in amplify-and-forward cooperative wireless networks with multiple source nodes," *EURASIP J. Wireless Commun. Netw.*, no. 1, pp. 1–13, 2012.
- [18] E. Beres and R. Adve, "Selection cooperation in multi-source cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 1, pp. 118–127, 2008.
- [19] C. H. Yu, B. Mumey, and O. Tirkkonen, "Distributed multiple relay selection by an auction mechanism," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, 2012.
- [20] S. Zhou, J. Xu, and Z. Niu, "Interference-aware relay selection scheme for two-hop relay networks with multiple source-destination pairs," *IEEE Trans. Veh. Technol.*, no. 99, 2013.

- [21] D. Li, Y. Xu, and J. Liu, "Distributed relay selection over multi-source and multi-relay wireless cooperative networks with selfish nodes," *Computer Communications*, vol. 33, no. 17, pp. 2145–2153, 2010.
- [22] B. Wang, Z. Han, and K. Liu, "Distributed relay selection and power control for multiuser cooperative communication networks using stack-elberg game," *IEEE Trans. Mobile Comput.*, vol. 8, no. 7, pp. 975–990, 2009.
- [23] A. Kurve, "Multi-user MIMO systems: the future in the making," *IEEE Potentials*, vol. 28, no. 6, pp. 37–42, 2009.
- [24] P. Bhat, S. Nagata, L. Campoy, I. Berberana, T. Derham, G. Liu, X. Shen, P. Zong, and J. Yang, "LTE-advanced: an operator perspective," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 104–114, 2012.
- [25] L. Yu, J. Liu, and K. Long, "Analysis of virtual MIMO-based cooperative communication in femtocell networks," in *IEEE Global Telecommun. Conf. (GLOBECOM)*, 2011, pp. 1–5.
- [26] M. Hong, Z. Xu, M. Razaviyayn, and Z.-Q. Luo, "Joint user grouping and linear virtual beamforming: Complexity, algorithms and approximation bounds," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2013–2027, 2013.
- [27] Y. Rui, H. Hu, H. Yi, and H. H. Chen, "Robust user pairing algorithm under channel estimation errors for uplink virtual multiple-input multiple-output systems," *IET Communications*, vol. 6, no. 3, pp. 318–323, 2012.
- [28] O. B. Karimi, M. A. Toutounchian, J. Liu, and C. Wang, "Lightweight user grouping with flexible degrees of freedom in virtual MIMO," *IEEE J. Sel. Areas Commun.*, 2013. [Online]. Available: <http://www.sfu.ca/~oba2/dls/jsac2013.pdf>
- [29] M. Haenggi, J. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, 2009.
- [30] J. Ilow and D. Hatzinakos, "Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers," *IEEE Trans. Signal Process.*, vol. 46, no. 6, pp. 1601–1611, Jun. 1998.
- [31] C. Buratti and A. Zanella, "Multihop virtual MIMO systems with channel reuse in a Poisson field of nodes," *IEEE Trans. Veh. Technol.*, vol. 60, no. 5, pp. 2060–2069, 2011.
- [32] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.
- [33] A. M. C. So, J. Zhang, and Y. Ye, "On approximating complex quadratic optimization problems via semidefinite programming relaxations," *Mathematical Programming*, vol. 110, no. 1, 2007.
- [34] CVX Research, Inc., "CVX: Matlab software for disciplined convex programming, version 2.0 beta," <http://cvxr.com/cvx>, Sep. 2012.
- [35] Z. Q. Luo, W. K. Ma, A. M. C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
- [36] T. Muetze, P. Stuedi, F. Kuhn, and G. Alonso, "Understanding radio irregularity in wireless networks," in *SECON*, 2008, pp. 82–90.
- [37] T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson, *Introduction to Algorithms*, 3rd ed. McGraw-Hill Higher Education, 2009.
- [38] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *Science*, vol. 315, no. 5814, pp. 972–976, 2007.
- [39] 3rd Generation Partnership Project (3GPP), "Further advancements for E-UTRA physical layer aspects (TR36.814, Release 9)," Tech. Rep., 2010. [Online]. Available: www.3gpp.org/Specs/htm-info/36814.htm
- [40] R. G. Gallager, *Discrete Stochastic Processes*. Kluwer, 1996.
- [41] B. Shah, "On the distribution of the sum of independent integer valued random variables," *American Statistician*, vol. 27, pp. 123–124, 1973.

PLACE
PHOTO
HERE

S. Hossein Seyedmehdi received the B.Sc. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran in 2005 and the M.Eng. degree in electrical and computer engineering from National University of Singapore, Singapore, in 2008. He is currently pursuing the Ph.D. degree in electrical and computer engineering at University of Toronto, Toronto, Canada.

In the summers of 2012 and 2013, he was a Research Intern at Ericsson, Ottawa, Canada where he worked on Smart Applications on Virtual Infrastructures (SAVI) project. His research interests include Information Theory for Wireless Communications, Signal Processing for MIMO channels, Algorithms for Radio Access Networks, and Future Radio Technologies.

PLACE
PHOTO
HERE

Gary Boudreau (M'84-SM'11) received a B.A.Sc. in Electrical Engineering from the University of Ottawa in 1983, an M.A.Sc. in Electrical Engineering from Queens University in 1984 and a Ph.D. in electrical engineering from Carleton University in 1989.

From 1984 to 1989 he was employed as a communications systems engineer with Canadian Astronautics Limited after which from 1990 to 1993 he worked as a satellite systems engineer for MPR Teltech Ltd. For the period spanning 1993 to 2009 he was employed by Nortel Networks in a variety of communication systems and management roles within the CDMA and LTE basestation product groups. In 2010 he joined Ericsson Canada where he is currently working in the LTE systems architecture group. His interests include digital and wireless communications as well as digital signal processing.