On the Performance of Reconfigurable Distributed MIMO in Mobile Networks

Zhong Zheng, *Member, IEEE*, and Zygmunt J. Haas, *Fellow, IEEE*

***Abstract*— We propose a new distributed Multi-Input Multi-Output (MIMO) architecture for mobile networks, which we refer to as reconfigurable distributed MIMO (RD-MIMO), where the communicating mobile nodes temporarily recruit adjacent nodes to operate as distributed antenna arrays. To best serve the communicating nodes, the node clusters are continuously reconfigured due to the node mobility and varying channel conditions. The frequency of reconfiguration depends on the required system performance, exhibiting a tradeoff between performance and complexity. We propose a practical node selection scheme, which activates only a small subset of transmitters. We evaluate the asymptotic performance of the scheme as a function of the number of recruited nodes, demonstrating that there is an optimal number of such nodes. Compared with the system that blindly activates all available transmitting nodes, our results show that the proposed RD-MIMO architecture with node selection achieves superior performance, especially as evident at low SNR. Furthermore, assuming the Brownian motion model, an analytical expression for the reconfiguration time to select a new cluster of transmitting nodes is obtained. Numerical results show that the obtained expression serves as a good estimation for the order of magnitude of the cluster reconfiguration time for other mobility patterns, such as the random walk model.**

***Index Terms*— Distributed MIMO, mobile networks, random matrix theory, Brownian motion.**

I. Introduction

A. MIMO and Distributed MIMO Technologies

IN MOBILE networks, the performance of wireless communications is limited by several channel impairments, such as multipath fading and distance-dependent path loss. To combat these effects, the spatial diversity techniques, such as Multi-Input Multi-Output (MIMO), have been used as means to improve the link level performance [1], [2]. For instance, in the traditional Co-located MIMO (C-MIMO) systems, where the antenna elements are implemented on single transmitter and/or receiver nodes, if the fading channels are uncorrelated, the link capacity scales linearly with the minimum number of transmit and receive antennas. However, when C-MIMO is applied to individual mobile devices, due to the physical dimensions of the devices, the distances between antenna elements on each transceiver are small compared to the wavelength of the carrier frequency, and the fading channels are usually correlated [3]. In addition, as the possible array size is small, the number of effective scattering objects is limited [4], [5] and the signal propagation is distorted by the so-called double scattering fading. Compared to the uncorrelated counterpart, the performance of the C-MIMO channels degrades, often significantly, under the impairments of antenna correlation [6], [7] and the double-scattering fading [8]–[10].

Recently, Distributed MIMO (D-MIMO) architecture has attracted increasing attention, where multiple transmit and receive stationary nodes cooperate to form distributed antenna arrays. In typical wireless networks, the communicating nodes are separated by much larger distance compared to the wavelength of the carrier frequency. At such distances, a large number of uncorrelated nodes can be grouped into the transmit and receive clusters [11], where each individual node may have only small number of antennas [12] or may see insufficient scattering environment [13]. Such formed D-MIMO systems can achieve capacity scaling similar to the uncorrelated C-MIMO systems [14], [15].

Previous works on the D-MIMO mainly focus on the cellular systems using remote antenna ports or base stations, which are connected to the central base station via high-speed fiber links. The purpose of such systems, termed Distributed Antenna System (DAS) [12] or Coordinated MultiPoint transmission (CoMP) [16], is to improve the coverage area and the spectral efficiency of the overall cellular systems, where the users' data streams are conveyed through the infrastructure networks. Therein, the performance analysis of the D-MIMO usually assumes a particular antenna placement with fixed locations [17]–[24] or stochastic antenna placements over a restricted topology, e.g., Wyner's circular model [11], [25]–[29]. These assumptions are relevant to the cellular networks, where the locations of the antenna ports are either fixed or can be controlled. Specifically, the asymptotic sum rate of large DAS networks with multiple noncooperative users was given in [11] using the Girko's law. The obtained expressions require solving fixed-point equations, and their solution does not have explicit form. By assuming a Wyner's circular antenna placement, the sum rate of DAS with multiple users were derived in [11], [26],and [25]. In the presence of composite fast fading and shadow fading, the performance of DAS was studied in terms of the ergodic capacity [13], [17]–[20], [29], [30] and outage capacity [20], [23], [30].

Manuscript received November 25, 2016; accepted January 13, 2017. Date of publication January 20, 2017; date of current version April 14, 2017. This work was supported by the NSF under Grant ECCS-1308208 and Grant ECCS-1533282. Part of this work will be presented at the 2017 Conference on Information Sciences and Systems. The associate editor coordinating the review of this paper and approving it for publication was V. Raghavan.

Z. Zheng is with the Department of Computer Science, University of Texas at Dallas, Richardson, TX 75080 USA (e-mail: zheng.zhong1@utdallas.edu).

Z. J. Haas is with the Department of Computer Science, University of Texas at Dallas, Richardson, TX 75080 USA, and also with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: zhaas@cornell.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCOMM.2017.2656129

Results in [22] and [31] show that the D-MIMO exhibits higher diversity and multiplexing gains, compared to the C-MIMO counterpart.

B. Reconfigurable-Distributed MIMO Architecture

To eliminate some of the above shortcomings of implementing MIMO systems on small footprint mobile devices, we propose a novel distributed MIMO architecture, where the mobile nodes can roam within a limited geographical area. The architecture is based on two relatively-distant nodes (referred to here as *cluster heads*), which desire to communicate over a channel with insufficient capacity. The cluster heads are allowed to *recruit* other close-by mobile nodes, as to form *transmit* and *receive clusters*, and the clusters form together a D-MIMO system. Importantly, as the nodes that create the MIMO clusters can move, the clusters need to be continually reconfigured to achieve a pre-defined degree of improvement. Due to these features, we refer to the proposed architecture as Reconfigurable Distributed MIMO (RD-MIMO).

The RD-MIMO scheme is applicable in Device-to-Device (D2D) communication scenario [32], as well as the next generation cellular networks, such as 5G systems and beyond, where a large number of mobile devices are available in relatively small areas [33], e.g., rooms, buildings, or microcells. In addition, modern cellular devices are typically equipped with multiple radio interfaces, including the primary cellular connection and direct connections among proximity devices, such as WiFi or WiFi Direct [34]. Therefore, when two nodes communicate with each other over a relatively large distance, each of them can temporarily recruit (using the direct high-capacity local WiFi links, as an example) a group of adjacent nodes to form the transmit and the receive clusters, so as to improve the long-distance communication. The transmissions between nodes in different clusters are carried out via the primary cellular radio interface, while the cooperations within each cluster are via the wireless high-capacity local links. Using the reliable communication among the cluster nodes [15], [35], the transmit/receive symbols at the cooperating nodes are jointly encoded/decoded, and the end-to-end equivalent channel behaves as a D-MIMO channel, approaching C-MIMO. As the RD-MIMO system is formed by mobile devices, the system architecture is subject to dynamic changes due to the node mobility. Timely cluster reconfiguration is required to fully utilize the benefit of the MIMO transmission.

Due to these distinct features of the RD-MIMO, we seek answers to the following questions: (1) how many cooperative nodes are needed to satisfy the performance requirements between communicating head nodes and (2) how often the head nodes should reselect new clusters due to varying node locations and link qualities. To address the first question, we consider a simple and practical transmitter selection scheme, which selects a subset of transmit nodes having the smallest path losses towards the receive nodes. The performance metric to determine the number of active transmit nodes is the average achievable rate using the node selection scheme, which is defined as the ergodic capacity of the RD-MIMO system, averaged over all possible realizations of the node placements. The average achievable rate has engineering intuition: the obtained rate is met for a typical node placement when nodes are placed randomly and independently [36]. This is relevant for the RD-MIMO systems, since the systems are formed by independent mobile nodes. As the systems are spanned over more diverse topologies, previous results on the capacity of D-MIMO systems with fixed node placement or stochastic Wyner's model cannot measure the system performance; thus, we derive new analytical expressions for the asymptotic average achievable rate using Random Matrix Theory (RMT).

To address the above second question, we consider a cluster reconfiguration scheme, which selects a new subset of transmit nodes periodically when the nodes move within the considered geographical area. The time interval between consecutive reconfigurations is selected statistically according to the node mobility pattern, where the active transmit nodes experience the smallest path losses towards the receive cluster with a given probability. Assuming that the transmit nodes are Brownian particles, we derive an approximation for the distribution of the time interval between consecutive cluster reconfiguration events. This approximation enables setting the cluster reconfiguration time analytically. We summarize the contributions of this paper as follows:

- Using RMT, we derive analytical expressions for the asymptotic average achievable rate of RD-MIMO. The considered rate is the average of the ergodic channel capacity over all possible realizations of node locations, which is applicable in the RD-MIMO scenarios due to random locations of mobile devices.
- Based on the derived rate expressions, we show that an optimal number of active transmit nodes exists to maximize the average achievable rate. As a comparison, when all transmit nodes are activated blindly, by increasing the number of transmit nodes, the achievable rate of RD-MIMO saturates to a limit that is less than the optimal setting.
- Assuming the transmit nodes move according to Brownian motion model, we derive an approximation for the time distribution between consecutive node reselection events. Based on this approximate distribution, explicit node reselection time is obtained, which periodically selects a subset of transmit nodes having the smallest path losses for a given probability. The computational complexity of the node (re-)selection scheme is shown to scale as $O(K^2 \log K)$, where K is the number of transmit nodes.
- In addition, we compare the analytical result of the node reselection time with the simulation result, assuming the practical random walk mobility model. The comparisons show that the obtained expression provides a good estimate for the order of magnitude of the cluster reconfiguration time, when the transmit nodes move as random walkers. Numerical results on the cluster reconfiguration time also show that the RD-MIMO system is feasible and could be implemented in realistic communication scenarios.

The rest of the paper is organized as follows. In Section II, we introduce the signal model and the node mobility model.

In Section III, we outline the node selection and reselection schemes. In Section IV, the average achievable rates of the RD-MIMO systems are derived, and the node reselection time is analyzed. Relevant numerical results are also provided. In Section V, we conclude the findings of the paper. Proofs of key technical results are derived in the Appendices.

II. System Model

Consider a wireless mobile communication network with K transmit and N receive nodes, where each transceiver is equipped with a single antenna. Among the transmit/receive nodes, there exists a transmit/receive cluster head node that initiates/terminates the user's data stream. The K transmit nodes form a cooperative transmit cluster, where the transmit head node is responsible to encode the information symbols into transmit signals, distributes the encoded signals to the corresponding assisting nodes, and synchronizes the transmissions within the cluster. Similarly, the N receive nodes form a cooperative receive cluster, where the receive head node collects the received signals from its assisting nodes and decodes the receive symbols. In this work, we assume that the node cooperation within each cluster is performed over reliable and high-speed local wireless connections, such as WiFi, while the transmissions between clusters are carried out via the separate cellular radio interface, prone to propagation impairments, such as fast and slow fading. As the local WiFi links have much higher rate compared to the cellular transmissions between remote clusters, the cellular transmissions are the bottleneck of the system, and the channels between the clusters resemble the distributed MIMO channels.

A. Signal Model

Given a transmit vector[1] $\mathbf{x}(t) = [x_1(t), \ldots, x_K(t)]^{\mathrm{T}}$, where $x_k(t)$ denotes the transmit signal of the node k at time t, the receive vector $\mathbf{y}(t) = [y_1(t), \ldots, y_N(t)]^{\mathrm{T}}$ is expressed as:

$$\mathbf{y}(t) = (\mathbf{G}(t) \circ \mathbf{H}(t))\,\mathbf{x}(t) + \mathbf{n}(t), \tag{1}$$

where $y_n(t)$ is the received signal at node n in the receive cluster and the operation $\circ$ denotes the entry-wise matrix multiplication. The fast fading coefficients between clusters are denoted as an $N \times K$ matrix $\mathbf{H}(t)$, where its entries are *i.i.d.* standard complex Gaussian random variables, with $h_{n,k}(t)$ denoting the coefficient of the block fading channel between the k^{th} transmit node and the n^{th} receive node. Letting T_c as the channel coherence time and $M \cdot T_c$ as the length of the coding block, the channel $\mathbf{H}(t)$ remains constant over each coherence period, and is *i.i.d.* across different coherence periods. The distance-dependent path losses are modeled as the matrix $\mathbf{G}(t)$, with the entry $g_{n,k}(t)$ denoting the square root of the average channel gain between the k^{th} transmit node and the n^{th} receive node, such that:

$$g_{n,k}(t) = c_p\, d_{n,k}(t)^{-\alpha/2}, \tag{2}$$

where c_p depends on the carrier frequency and represents the square root of path loss at unit distance, α (≥ 2) is the path-loss exponent, and $d_{n,k}(t)$ is the distance between the k^{th} transmit node and the n^{th} receive node at time t. Other large-scale fading, such as shadowing, is ignored in the signal model (1). The additive noise $\mathbf{n}(t)$ is modeled as an *i.i.d.* complex Gaussian vector with power ρ^2, i.e., $\mathbf{n}(t) \sim \mathcal{CN}(\mathbf{0}, \rho^2\mathbf{I})$. In this work, we have adopted the following assumptions for the signal model:

(A1) Within each cluster, the head node performs joint encoding/decoding of the transmitted/received signals. With relatively low node mobility, the path loss $\mathbf{G}(t)$ varies slowly over the duration of multiple coding blocks. For each block interval $M \cdot T_c$, $\mathbf{G}(t)$ is treated as a constant matrix, and the fast fading $\mathbf{H}(t)$ is ergodic.

(A2) In the transmit cluster, the jointly encoded signal $\mathbf{x}(t)$ is perfectly conveyed from the head node to the assisting nodes via the local WiFi links. In the receive cluster, the received signal $\mathbf{y}(t)$ and the CSIs are perfectly conveyed from the assisting nodes to the head node. This is justified by the fact that the local WiFi links are assumed to have much larger rate than the cellular links.

(A3) At the transmit nodes, the transmit signal $\mathbf{x}(t)$ is Gaussian distributed with covariance matrix[2] $\mathbb{E}[\mathbf{x}(t)\mathbf{x}(t)^{\dagger}] = \mathbf{Q}$, i.e., $\mathbf{x}(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{Q})$.

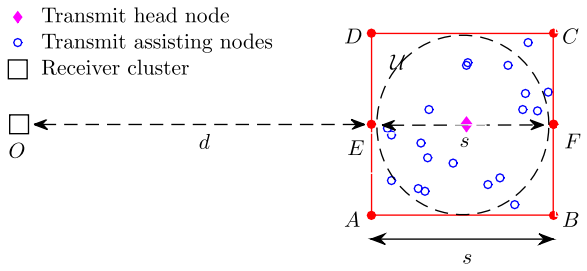


Fig. 1. RD-MIMO with two cooperative clusters with inter-cluster distance d. The edge length of the square transmit cluster is s. A circular transmit node placement is shown as a dashed circle.

In Fig. 1, we show an example of a transmit and a receive cluster, spanned over a planar area. Typically, the shape of the transmit cluster will practically be assumed to be a circle (assuming omni-directional antennas), and could thus be modeled as the dashed circle in Fig. 1. This cluster configuration corresponds to the case when the cluster is formed by the nodes within certain distance to the head node. The distance between the closest edges of the two clusters is denoted as d, with the receive cluster located at the origin O. The transmit nodes are randomly distributed in the circular area $\mathcal{U}$ of radius $s/2$ and centered at the transmit head node. However, it turns out that a square model (as shown in Fig. 1 in red) allows further analytical treatments for the achievable rate and the cluster reselection times of the RD-MIMO system. Therefore, we adopt the square model as an approximation to the above circular model, and we demonstrate in Section IV that the "error" introduced by the square model (as opposed to the circular model) is negligible. The analogous square model assumes that the cluster area $\mathcal{U}$ is a square, with the edge

[1] $(\cdot)^{\mathrm{T}}$ denotes the transpose operation.

[2] $(\cdot)^{\dagger}$ denotes complex conjugate and transpose operation.

length of $\mathcal{U}$ denoted as s. To obtain tractable analysis and to gain insights into the impact of the distributed transmit nodes, we assume that the distance d is much larger than the dimension of the receive cluster. In this setting, the average channel gain is approximated as:

$$g_{n,k}(t)^2 \approx g_k(t)^2, \quad 1 \le n \le N \text{ and } 1 \le k \le K. \tag{3}$$

In the following sections, our analysis is based on the square cluster model, and the simulations in Section IV show that the obtained analytical results are good estimate of the performance of RD-MIMO with the circular node placement.

B. Mobility Model

Denote the locations of the transmit nodes at time t as $a_1(t), \ldots, a_K(t)$, where $0 \le t \le t_{\max}$, and $t_{\max}$ is the stopping time. The location of the i^{th} transmit node is represented as $a_i(t) = p_i(t) + \mathrm{j}\, q_i(t)$, where $p_i(t)$ and $q_i(t)$ are coordinates on the Euclidean plane. We model a node motion along the two coordinates as independent Brownian motions with reflected boundaries as described below. We consider that the transmit nodes move within the square cluster area $\mathcal{U}$, as shown in Fig. 1, i.e., $d \le p_i(t) \le d+s$ and $-s/2 \le q_i(t) \le s/2$ for $0 \le t \le t_{\max}$. Define K *i.i.d.* Brownian motions with diffusion coefficient D over the complex plane as $\mathbf{W}_i = \{w_i(t) = u_i(t) + \mathrm{j}v_i(t),\ 0 \le t \le t_{\max}\}$, $1 \le i \le K$, where $u_i(t)$ and $v_i(t)$ are the coordinates on the Euclidean plane, and $u_i(t+h) - u_i(t)$ and $v_i(t+h) - v_i(t)$ are independent Gaussian random variables with zero mean and variance $Dh/2$. Intuitively, a diffusion coefficient D (in squared meter per second) is the mean squared displacement of Brownian motion over one second, i.e., $\mathbb{E}[|w_i(t+1) - w_i(t)|^2] = D$. We assume that the mobility of the transmit nodes is modeled as *i.i.d.* Brownian motions $\{\mathbf{W}_i\}_{1 \le i \le K}$ with reflected boundaries of $\mathcal{U}$, such that:

$$p_i(t) = \begin{cases} u_i(t), & t \in \mathcal{T}_{i,1} \\ 2d - u_i(t), & t \in \mathcal{T}_{i,2} \\ 2(d+s) - u_i(t), & t \in \mathcal{T}_{i,3} \end{cases},$$

$$q_i(t) = \begin{cases} v_i(t), & t \in \mathcal{S}_{i,1} \\ -s - v_i(t), & t \in \mathcal{S}_{i,2} \\ s - v_i(t), & t \in \mathcal{S}_{i,3} \end{cases}, \tag{4}$$

where $t \in \mathcal{T}_{i,1}$ when $d < u_i(t) < d+s$, $t \in \mathcal{T}_{i,2}$ when $u_i(t) \le d$, $t \in \mathcal{T}_{i,3}$ when $u_i(t) \ge d+s$, $t \in \mathcal{S}_{i,1}$ when $-s/2 < v_i(t) < s/2$, $t \in \mathcal{S}_{i,2}$ when $v_i(t) \le -s/2$, and $t \in \mathcal{S}_{i,3}$ when $v_i(t) \ge s/2$.

III. Transmit Node Selection and Cluster Reconfiguration

We consider practical node selection and cluster reconfiguration schemes, which only utilize the knowledge of path losses. The considered schemes activate a subset of transmit nodes for inter-cluster transmission and periodically reconfigure the transmit cluster to best serve the communicating head nodes, while reducing the operational costs, such as the signaling overhead to form the transmit cluster. In this section, we describe the procedures for the node selection and the cluster reconfigurations, as well as the key performance metrics to configure these schemes, such as the average achievable rate and the cluster reselection time. The performance of the proposed schemes and the performance metrics are quantitatively evaluated in Section IV.

A. Transmit Node Selection

Using channel sounding techniques similar to what is done in LTE [37], the transmit nodes can measure the path losses of their own channels to the receive cluster nodes. The measurements are then conveyed to the transmit head node using the local links. Utilizing the path loss matrix $\mathbf{G}(t)$, a subset of L nodes (out of the K nodes) in the transmit cluster is activated for the RD-MIMO transmission; these L nodes are selected based on their smallest path losses towards the receive cluster. The L active transmit nodes are assumed to track the variation of path losses for rate adaptation purposes.[3] When the number of available transmit nodes is large, selecting a small subset of nodes could reduce the power consumption and signaling overheads induced by the channel measurements and measurement feedback. In addition, as will be shown in Section IV, the RD-MIMO channel may achieve an optimal average rate by properly selecting the number of active nodes.[4] A possible node selection procedure is detailed below:

(P1) The head node in the transmit cluster determines the size of $\mathcal{U}$ by system considerations, such as the coverage area of wireless local links, the node density, the fading characteristics, etc. Nodes in $\mathcal{U}$ are reachable by the head node transmitting with certain power on the local links. Assume that there are K nodes in $\mathcal{U}$.

(P2) The head node transmits a "beacon message" to the nodes in the area $\mathcal{U}$ triggering the K candidate transmit nodes[5] to initiate channel measurements of the path losses of their own channels towards the receive cluster. The channel measurements are then sent to the transmit head node by local links. The channel measurements can be determined either via the inverse link using channel reciprocity or through feedback from the receive nodes.

(P3) The head node activates L transmit nodes ($L \le K$) having the largest average channel gains, by sending the nodes an "activation message", as to form the transmit cluster, i.e., the average channel gains of the selected active nodes are:

$$g_{(1)}(t) \ge g_{(2)}(t) \ge \ldots \ge g_{(L)}(t), \tag{5}$$

where $g_{(i)}(t)$ is the i^{th} largest channel gain among $\{g_1(t), \ldots, g_K(t)\}$ at time t. The transmit power is equally allocated to the active transmit nodes, while inactive nodes remain muted.

[3] Again, this could be done, for example, using a similar scheme to channel sounding in LTE, although other implementations are possible as well. Due to space limitations, we omit further implementation discussion here.

[4] As is discussed later, more nodes do not, in general, lead to better performance; there is a target number of nodes that should be selected to optimize the performance.

[5] For reducing complexity of the scheme, a smaller number of nodes may be triggered to perform the measurements, especially when the node density is large. See Section IV-B for a detailed discussion.

Note that, using the knowledge of path losses, power optimization across transmit nodes can be performed to further improve the achievable rate. However, in most of the mobile networks, the transmit power of each node is managed by the node itself and is subject to hardware limitations of the devices. Thus, we assume here equal power for the active transmit nodes, and we leave the issue of power optimization with per-node constraint for future work. A MIMO technique analogous to the node selection scheme is the Transmit Antenna Selection (TAS) in C-MIMO, which chooses a subset of antenna elements depending on the fast fading channel coefficients. Among multi-antenna techniques, TAS is attractive, since it significantly reduces the hardware implementation costs [38]. The TAS scheme has been also motivated in multiuser scheduling by [39] due to a higher multiuser diversity gain compared to other MIMO systems. However, our proposed node selection scheme is different from TAS, since the selection criterion (5) depends on the path loss measurements, which results in different system behaviors and performance compared to the TAS schemes.

Due to the approximation (3) and the selection criterion (5), the received signal $\mathbf{y}(t)$ in (1) after transmit node selection can be rewritten as:

$$\mathbf{y}_{(L)}(t) = \mathbf{H}_{(L)}(t)\,\mathbf{G}_{(L)}(t)\,\mathbf{x}_{(L)}(t) + \mathbf{n}(t), \tag{6}$$

where $N \times L$ matrix $\mathbf{H}_{(L)}(t)$ denotes the fast fading coefficients between the active transmitters and the receive cluster, and $\mathbf{G}_{(L)}(t)$ is a diagonal matrix with the i^{th} diagonal entry being $g_{(i)}(t)$. The entry-wise matrix multiplication in (1) reduces now to the standard matrix multiplication in (6). Due to the equal power allocation at the active transmit nodes, the transmitted signal is distributed as $\mathbf{x}_{(L)}(t) \sim \mathcal{CN}(0, (P/L)\,\mathbf{I})$, where P is the total transmit power of the L transmitting nodes. Given the assumptions (AII-A)–(AII-A) in Section II-A, the instantaneous capacity of the channel (6) is given by [11] in nats/s/Hz as[6]:

$$\begin{aligned} C_L &= \frac{1}{N}\log\det\left(\mathbf{I} + \frac{\gamma}{L}\mathbf{H}_{(L)}\Sigma_{(L)}\mathbf{H}_{(L)}^{\dagger}\right) \\ &= \frac{1}{N}\sum_{i=1}^{N}\log\left(1 + \frac{\gamma}{\zeta}\lambda_i\right), \end{aligned} \tag{7}$$

where the time index t is dropped wherever it is clear from the context. The matrix $\Sigma_{(L)} = 1/g_0^2\mathbf{G}_{(L)}^2$ is diagonal with the i^{th} diagonal entry being $\sigma_{(i)} = (g_{(i)}/g_0)^{\alpha} = (d/d_{(i)})^{\alpha}$, where $d_{(1)} \le \cdots \le d_{(K)}$ are the ordered distances between the transmit nodes and the receive cluster, within which the receive nodes have the same location due to the approximation (3). We denote $g_0 = c_p d^{-\alpha/2}$ and $\gamma = Pg_0^2/\rho^2$ as the path loss and the receive SNR, respectively, as if the transceivers are located at O and E. We denote λ_i, $1 \le i \le N$, as the eigenvalues of the random matrix $\mathbf{H}_{(L)}\Sigma_{(L)}\mathbf{H}_{(L)}^{\dagger}/N$ and $\zeta = L/N$.

Clearly, the performance of RD-MIMO depends on the number of active transmit nodes. It is thus important to choose a proper value of L to optimize the node selection. Herein, the performance of the node selection is evaluated via the average achievable rate of the RD-MIMO system, defined as $\mathcal{R}_L = \mathbb{E}[C_L]$, where the expectation is taken over all realizations of path loss $\mathbf{G}_{(L)}$ and fast fading $\mathbf{H}_{(L)}$. The average achievable rate has engineering intuition: it is the expected rate for a typical node placement when each node is placed randomly and independently.

[6]$\det(\cdot)$ denotes the determinant of matrix.

B. Cluster Reconfiguration

As the transmit nodes move randomly within the cluster area, the subset of nodes with the largest channel gains towards the receive cluster changes over time. When some inactive nodes experience more favorable channel conditions compared to the active nodes, reselection of a new subset of transmit nodes may be needed to improve the performance of RD-MIMO. However, since the path losses are tracked and fed back to the head node only by the active transmitters as to reduce signaling overhead,[7] a scheme is needed to initiate the channel feedback from inactive nodes and for the node reselection. To timely capture the reselection event, the transmit head node periodically triggers the channel measurements and feedback from all the candidate transmit nodes to the head node. The frequency of the cluster reselection needs to be properly set to achieve a tradeoff between the system performance and complexity, e.g., operational costs. In other words, although a frequent reconfiguration would ensure good system performance, it may come at the expense of prohibitively large complexity of the system operation. On the other hand, an insufficiently frequent reconfiguration would every-so-often leave the system in too suboptimal condition. In what follows, we evaluate this tradeoff for a particular system setting, where the reselection time is determined statistically according to the node mobility pattern. Note that when the capacity of the RD-MIMO channel drops below a desired threshold, the transmit head node may also trigger the reselection before the periodic reselection time is reached.

Without loss of generality, the initial distances between the receive cluster $\big($at the origin $(0,0)\big)$ and the transmit nodes are ordered as $|a_1(0)| \le \cdots \le |a_K(0)|$, and the nodes with indexes $1, \ldots, L$ are the selected active transmit nodes. Due to the criterion (5) and $g_i(t)$ being a decreasing function of the distance, the node reselection needs to be performed when any of the active node has larger distance compared to any of the inactive node. Given the initial node locations, the time instance to perform a node reselection is defined as the first hitting time τ between the active and the inactive nodes, such that:

$$\tau = \inf\left\{t \ge 0 : \max_{1\le i\le L}|a_i(t)| > \min_{L+1\le i\le K}|a_i(t)|\right\}. \tag{8}$$

To trade off between timely reconfigurations and system complexity, the periodic reselection interval is chosen to capture a fraction of reselection events as defined in (8). Specifically, for a fixed ϵ $(0 < \epsilon < 1)$, the reselection interval t_ϵ satisfies:

$$\epsilon = \Pr(\tau < t_\epsilon) = H(t_\epsilon), \tag{9}$$

[7]In fact, even the active nodes may limit the path losses tracking to reduce complexity of the scheme.

where $H(\cdot)$ denotes the distribution function of the first hitting time τ when the transmit nodes move according to the Brownian motion model as defined in Section II-B. The reselection interval t_ϵ is set to capture $(1-\epsilon)$ of all reselection events as defined by the reconfiguration criteria. For instance, if (8) defines a reconfiguration criteria, then $t_{0.1}$ will indicate the time duration in which (on the average) at least 90% of the reconfiguration events (8) occur. Said differently, if we periodically reconfigure the system every $t_{0.1}$, then we expect that (on the average) the active nodes have the smallest path losses towards the receive cluster in 90% of the time. By decreasing ϵ, the number of initiated cluster reselection operations is increased as to maintain good performance at the expense of more frequent channel measurements and feedback from the inactive transmit nodes.

IV. Performance Analysis of RD-MIMO

In this section, we derive asymptotic expressions for the average achievable rate using tools for the spectral analysis of large dimensional random matrices. Assuming transmit nodes are Brownian particles, the cluster reselection time is obtained for the RD-MIMO systems. Based on the analytical expressions, we answer the following questions regarding the (re-)configuration of a transmit cluster: (1) how many active nodes are needed, and (2) how often a reselection should take place. Numerical simulations in this section are conducted with MATLAB R2015a.

A. Asymptotic Average Achievable Rate

To gain insight into the performance of RD-MIMO, we adopt an asymptotic analysis for the spectrum of large random matrix $\mathbf{H}_{(L)}\Sigma_{(L)}\mathbf{H}_{(L)}^{\dagger}/N$. In the following, we consider the asymptotic regime of:

$$L,\ N\to\infty,\quad \text{with } 0<\zeta=\frac{L}{N}<\infty. \tag{10}$$

Numerical results in this section show that the obtained asymptotic results, assuming (10), serve as good approximation for a practical RD-MIMO system with "not-so-large" L and N. Similar observations have been also pointed out by [5], [7], [8],and [10] for other MIMO configurations.

In the following proposition, the average achievable rate of RD-MIMO is given in the asymptotic regime (10) with integer value of the path-loss exponent α.

Proposition 1: In the asymptotic regime (10) with integer values of α, the average achievable rate $\mathcal{R}_L=\mathbb{E}[C_L]$ almost surely converges to a nonrandom limit as:

$$\mathcal{R}_L \xrightarrow{(10)} \zeta\,\mathcal{V}\left(\frac{\eta\,\gamma}{\zeta}\right)+\log\frac{1}{\eta}+\eta-1, \tag{11}$$

where $\eta\geq 0$ is the solution of the fixed-point equation:

$$\zeta=\frac{1-\eta}{1-\zeta/(\gamma\ \eta)\ \mathcal{G}\big(-\zeta/(\gamma\ \eta)\big)}. \tag{12}$$

The functions $\mathcal{V}(\gamma)$ and $\mathcal{G}(\omega)$ are given in (13) and (14) on the bottom of this page page, where $\theta=d/s$ and ${}_3F_2$ denotes the generalized hypergeometric function [40, eq. (9.14.1)].

Proof: The proof of Proposition 1 is a direct application of [41, Th. 2.39] using the spectral distribution of $\Sigma_{(L)}$. The detailed proof is provided in Appendix A. ∎

We note that the intermediate result [41, Th. 2.39] was originally used to calculate the asymptotic capacity of the C-MIMO channels in presence of spatial correlations, where $\Sigma_{(L)}$ is the correlation among the transmit antennas. This research direction has also been pursued in [42]–[46] and the references therein, where the obtained performance metrics usually depend on certain realization of $\Sigma_{(L)}$ or the empirical distribution of its entries relevant in the context of antenna correlation. However, the average achievable rate obtained in Proposition 1 requires averaging over *all* realizations of path losses $\Sigma_{(L)}$ between clusters, which requires non-trivial follow-up derivations from [41, Th. 2.39]. The obtained result enables us to make relevant observations on the behaviors of RD-MIMO systems, as will be discussed in Corollaries 1-3.

The asymptotic rate $\mathcal{R}_L$ obtained by (11) requires one to evaluate the fixed-point equation (12). In general cases, the solution is intractable, and does not yield insight into the behavior of the RD-MIMO systems. However, in the next subsections, we utilize Proposition 1 to obtain explicit expressions of $\mathcal{R}_L$ for high and low SNR regimes. Based on the explicit expressions, we observe that the average achievable rate can be maximized by optimally selecting L, which is typically smaller than the total number of the candidate nodes K. In addition, we obtain the explicit rate expression when the transmit head node blindly activates all transmit nodes. As the number of transmit nodes increases, we observe that the achievable rate saturates to a limit that is less than the rate achieved by the optimal node selection setting. The results on the asymptotic rate $\mathcal{R}_L$ and the optimal number of selected nodes L are vital for the operation of the cluster (re-)configuration schemes.

1) Optimal Number of Active Transmit Nodes: We assume next that the inter-cluster distance d is much larger than the edge length s such that $d\to\infty$ while s fixed. Based on Proposition 1, simple and intuitive rate expressions are obtained in Corollary 1 in high SNR ($\gamma\to\infty$) and low SNR ($\gamma\to 0$) regimes, respectively.

Corollary 1: Assume $d\to\infty$ while s fixed. The average achievable rate $\mathcal{R}_L$ at high SNR ($\gamma\to\infty$) is

$$\mathcal{V}(\gamma)=\frac{1}{\alpha}\sum_{i=1}^{\alpha}\int_0^{\gamma}\frac{x^{\frac{1}{\alpha}-1}}{x^{\frac{1}{\alpha}}-e^{\frac{2i+1}{\alpha}\pi\mathrm{j}}}\,{}_3F_2\left(1,1,L+1;2,K+1;\frac{1}{\theta(x^{\frac{1}{\alpha}}e^{-\frac{2i+1}{\alpha}\pi\mathrm{j}}-1)}\right)\mathrm{d}x, \tag{13}$$

$$\mathcal{G}(\omega)=\frac{1}{\alpha\,\omega}\sum_{i=1}^{\alpha}\frac{1}{1-|\omega|^{\frac{1}{\alpha}}e^{\frac{2i+1}{\alpha}\pi\mathrm{j}}}\,{}_3F_2\left(1,1,L+1;2,K+1;\frac{1}{\theta(|\omega|^{-\frac{1}{\alpha}}e^{-\frac{2i+1}{\alpha}\pi\mathrm{j}}-1)}\right)-\frac{1}{\omega}, \tag{14}$$

given by:

$$\mathcal{R}_L = \begin{cases} \zeta \log\left(\dfrac{\gamma}{\zeta}\right) + (\zeta - 1)\log(1-\zeta) - \dfrac{(L+1)\alpha\,\zeta}{2(K+1)\theta} - \zeta, \\ \qquad\qquad\qquad\qquad \zeta \le 1 \\ \log(\gamma) + (\zeta - 1)\log\left(\dfrac{\zeta}{\zeta - 1}\right) - \dfrac{(L+1)\alpha}{2(K+1)\theta} - 1, \\ \qquad\qquad\qquad\qquad \zeta > 1. \end{cases} \tag{15}$$

At low SNR ($\gamma \to 0$), the average achievable rate $\mathcal{R}_L$ is given by:

$$\mathcal{R}_L = \zeta \log\left(1 + \frac{\gamma}{\zeta}\right) - \frac{(L+1)\alpha\,\zeta}{2(K+1)\theta}\frac{\gamma}{\gamma + \zeta}. \tag{16}$$

Proof: The proof of Corollary 1 is in Appendix B. ■

According to Corollary 1, when the number of active transmit nodes L is large, $\mathcal{R}_L$ is a decreasing function of L. To see this, we note that when $\zeta \gg 1$, $(\zeta - 1)\log\left(\zeta/(\zeta - 1)\right) \approx 1$. Thus, the rate $\mathcal{R}_L$ in high SNR regime behaves as:

$$\mathcal{R}_L \sim \log\gamma - \frac{(L+1)\alpha}{2(K+1)\theta}, \quad \gamma \to \infty. \tag{17}$$

Similarly, in low SNR regime, we have $\zeta \log(1 + \gamma/\zeta) \approx \gamma$, and thus the rate $\mathcal{R}_L$ behaves as:

$$\mathcal{R}_L \sim \gamma\left(1 - \frac{(L+1)\alpha}{2(K+1)\theta}\right), \quad \gamma \to 0. \tag{18}$$

Both (17) and (18) show that $\mathcal{R}_L$ is a decreasing function of L, and that a higher path-loss exponent α leads to more severe rate degradation. In addition, the average achievable rate reduces more rapidly at low SNR, as $\mathcal{R}_L$ is proportional to $1 - (L+1)\alpha/(2(K+1)\theta)$. This is in contrast to the high SNR case, where the rate reduction $(L+1)\alpha/(2(K+1)\theta)$ is marginal compared with large value of $\log\gamma$. On the other hand, when L is small, the capacity C_L in (7) may be improved by activating additional transmit nodes, if the channel gains of the activated nodes exceed a certain threshold. For instance, by setting the channel gain $\sigma_{(1)} = 1$, the difference $\left(\exp(C_2) - \exp(C_1)\right)$ can be lower-bounded by using the inequality [47, eq. (18)] as[8]:

$$\exp(C_2) - \exp(C_1) \ge \left(1 + \frac{\gamma\sqrt{\sigma_{(2)}}}{2}\exp\frac{\psi(N) + \psi(N-1)}{2}\right)^2 - 1 - N\gamma, \tag{19}$$

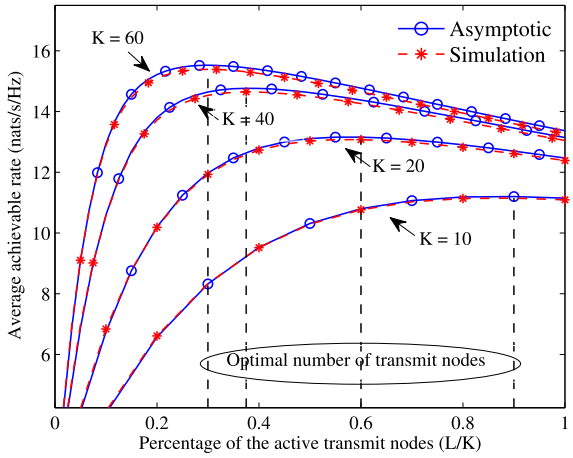
where $\psi(\cdot)$ denotes the digamma function [40, eq. (8.360/1)]. By expanding the first term in the right-hand-side (RHS) of (19) and ignoring the term with the smaller power of γ, at high SNR the capacity can be improved by including one additional transmit node when:

$$\sigma_{(2)} \ge \frac{4N}{\gamma\exp(\psi(N) + \psi(N-1))}. \tag{20}$$

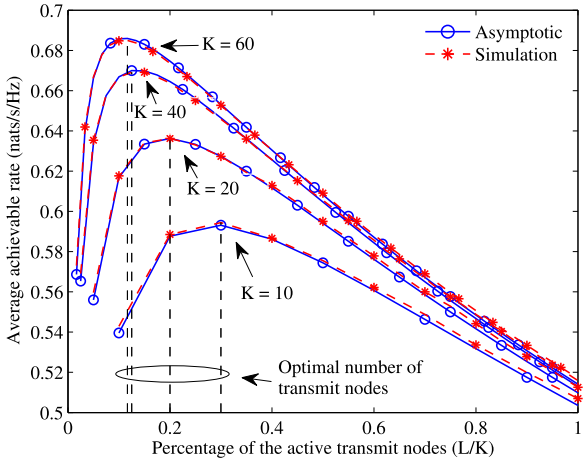
In addition, using approximation [18, eq. (12)], the ratio between C_2 and C_1 at low SNR can be lower-bounded by:

$$\frac{C_2}{C_1} \ge \frac{1+N}{1 + N\frac{\sigma_{(2)}^2 + 1}{(\sigma_{(2)}+1)^2}}\frac{\log(N\gamma\sqrt{\sigma_{(2)}})}{\log(N\gamma)}. \tag{21}$$

[8] C_1 and C_2 are C_L for $L = 1$ and 2, respectively.



(a)



(b)

Fig. 2. The average achievable rate of RD-MIMO with transmit node selection. The number of receive nodes is $N = 8$. Lines labeled with "o" denote average rates obtained by (11), lines labeled with "∗" denote simulation results, and dashed vertical lines denote the optimal percentage of active transmit nodes. (a) SNR $\gamma = 10$ dB (b) SNR $\gamma = -10$ dB.

To obtain a larger C_2 compared with C_1, the channel gain $\sigma_{(2)}$ should satisfy:

$$\frac{(\sigma_{(2)}+1)^2}{\sigma_{(2)}}\log\frac{1}{\sigma_{(2)}} \le \frac{4N\log(N\gamma)}{N+1}. \tag{22}$$

It is noted that the left-hand-side (LHS) of (22) is a decreasing function of $\sigma_{(2)}$ when $0 \le \sigma_{(2)} \le 1$ and becomes positive infinity as $\sigma_{(2)} \to 0$. Then, there exists a minimum $\sigma_{\min}$ that satisfies (22) with equality. Therefore, at low SNR, the capacity improvement by using one additional transmit node requires $\sigma_{(2)} \ge \sigma_{\min}$. To sum up, when the channel gain $\sigma_{(2)}$ fulfills the condition (20) or (22), the capacity C_L is increasing as a function of L at $L = 1$ and decreasing when L is comparable to K, i.e., $L \approx K$. Thus, there exists an optimal L that maximizes C_L. In addition, when SNR is small, the rate reduction is more severe at large L and it is more beneficial to select a smaller number of transmit nodes in this operational regime.

In Fig. 2, we investigate the impact of the number of active transmit nodes in the large and low SNR regimes.

Assuming that the number of receive nodes is fixed to be $N = 8$, the number of candidate transmit nodes is set to be $K = 10,\ 20,\ 40,$ and 60, and that the transmit nodes are uniformly distributed in the square area $\mathcal{U}$ with $s = 15$ meters. The distance between the transmit and the receive clusters is $d = 30$ meters with path-loss exponent $\alpha = 4$. Fig. 2(a) shows the average achievable rate (11) as a function of the percentage of active transmit nodes L/K, when SNR $\gamma = 10$ dB. As L increases, the average achievable rate drastically improves for a relatively small increase in L, and the performance varies only mildly when more transmit nodes becomes active (above 60% when $K = 10$ and 20, and above 30% when $K = 40$ and 60). Furthermore, as the ratio L/K approaches 1, the performance of RD-MIMO actually degrades for $K \geq 20$. Fig. 2(b) shows the average achievable rate when the SNR $\gamma = -10$ dB. In this case, the optimal rate can be achieved by utilizing only a small number of the transmit nodes, e.g., 3-6 best nodes from the available K transmitters. When L exceeds this optimal setting, the achievable rate decreases proportionally to L. These results indicate that an optimal rate can be achieved by using the proposed node selection scheme. This is in line with our predictions based on (17)–(22). The node selection scheme is especially effective in the low SNR regime, as the optimal rate can be attained by selecting a smaller subset of transmit nodes.

2) Rate Saturation With Blind Transmit Nodes: In this RD-MIMO scenario, we assume that the transmit head node blindly activates all candidate nodes in the cluster area $\mathcal{U}$, i.e., $L = K$. Based on Proposition 1, we show in the following corollaries that the resulting average achievable rate saturates to a limiting rate when the number of transmit nodes increases. On the other hand, numerical results show that for a given SNR, the fluctuation of the ergodic capacity diminishes and that the ergodic capacity is (nearly) independent of realizations of path losses between the transmit nodes and the receive cluster. Therefore, the rate adaptation of such RD-MIMO systems is easy to implement, where the transmit nodes do not need to track the path losses, and no cluster reconfigurations are required to cope with the changes in path losses.

Corollary 2: In the asymptotic regime (10) with $L = K$, the average achievable rate $\mathcal{R}_K$ is given by (11) with $\mathcal{V}(\cdot)$ and $\mathcal{G}(\cdot)$ reducing to:

$$\mathcal{V}(\gamma) = (\theta+1)\log\left(1+\left(\frac{\theta}{\theta+1}\right)^{\alpha}\gamma\right) - \theta\log(1+\gamma) + \alpha\mathcal{F}(-\gamma) + \alpha, \quad (23)$$

$$\mathcal{G}(\omega) = \frac{1}{\omega}\mathcal{F}\left(\frac{1}{\omega}\right), \quad (24)$$

where $\theta = d/s$. The function $\mathcal{F}(\cdot)$ is given by:

$$\mathcal{F}(x) = \theta\,{}_2F_1\left(1, -\frac{1}{\alpha}; 1-\frac{1}{\alpha}; x\right) - (\theta+1)\,{}_2F_1\left(1, -\frac{1}{\alpha}; 1-\frac{1}{\alpha}; \left(\frac{\theta}{\theta+1}\right)^{\alpha}x\right), \quad (25)$$

where ${}_2F_1$ denotes the Gauss hypergeometric function [40, eq. (9.100)].

Proof: The proof of Corollary 2 is in Appendix C. ■

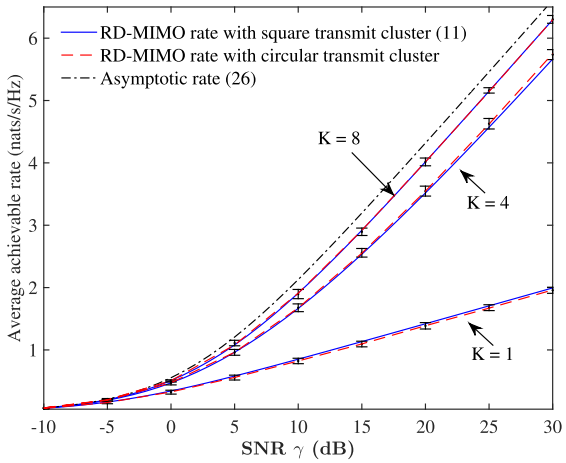


Fig. 3. Average achievable rate of RD-MIMO without node selection ($L = K$). The number of receive nodes is $N = 4$, inter-cluster distance is $d = 30$, the edge length of transmit cluster is $s = 5$, and the path-loss exponent is $\alpha = 4$. Solid lines: average rate obtained by Corollary 2; dash-dotted line: average rate obtained by (26); dashed lines: rate of RD-MIMO when the transmit nodes are distributed in a circular area as shown in Fig. 1; vertical lines: standard deviation of ergodic capacity C_K around average rate $\mathcal{R}_K$ calculated from 10^3 realizations of node placement.

Although Corollary 2 is derived assuming integer values of α, (23) and (24) are valid for any path-loss exponent α (≥ 2) being a real number. This can be verified by taking the expectations in (33) over the distribution of $\{\sigma_k\}_{1\leq k\leq K}$. We omit the derivations due to the page limitation.

Corollary 3: In the asymptotic regime (10) with $L = K$, the average achievable rate $\mathcal{R}_K$ for large but finite $\zeta = K/N$ is given by:

$$\mathcal{R}_K = \log\left(1+\frac{\theta}{\alpha-1}\left(1-\left(\frac{\theta}{\theta+1}\right)^{\alpha-1}\right)\gamma\right). \quad (26)$$

Proof: The proof of Corollary 3 is in Appendix C. ■

In the case of blind transmitter, Corollary 3 states that the average achievable rate of RD-MIMO approaches the limit (26) by increasing the number of transmit nodes K within the fixed cluster area. As the spatial degree of freedom is limited by the number of receive nodes N when $\zeta > 1$, the rate of RD-MIMO with blind transmit nodes cannot be efficiently improved by densifying distributed transmit antennas.

Fig. 3 shows the average achievable rate $\mathcal{R}_K$ assuming that there are $N = 4$ receive nodes and $K = 1,\ 4,$ and 8 transmit nodes. Using Corollary 2, we obtain the rate $\mathcal{R}_K$ with the inter-cluster distance $d = 30$ meters and with the edge length of transmit cluster $s = 5$ meters. The path-loss exponent α is set to 4. As the number of transmit nodes increases, the improvement of the average achievable rate is large when K is relatively small (from 1 to 4 transmit nodes), especially in the high SNR regime. However, when K is large, the improvement becomes marginal and the achievable rate of RD-MIMO approaches the limit given by (26), as predicted by Corollary 3. These results show that, in the case of blind transmitters, the RD-MIMO formed by a relatively small number of transmit nodes (e.g., $K = 8$ in this setting) can attain almost full rate achieved by densifying transmit

cluster with $K \to \infty$, while N fixed. When a large number of transmit nodes is available, only a small subset of nodes needs to be activated to tradeoff between the exploitable performance and overheads induced by forming the transmit antenna array. In addition, the standard deviations of the ergodic capacity C_K calculated from 10^3 realizations of node placements are plotted as vertical lines centered at the average rate $\mathcal{R}_K$. It is clear that the fluctuations of ergodic capacity due to the network topology is marginal as long as all candidate transmit nodes are activated. This fact further reduces the signaling overhead within the transmit cluster, since the path loss measurements from transmit nodes are not needed for rate adaptation at the head node. As a comparison, we also plot the the achievable rate of RD-MIMO when the transmit nodes are uniformly distributed in a circular area, as shown in Fig. 1. Results in Fig. 3 show that the Corollary 2 also accurately estimates the achievable rate in the circular cluster configuration.

B. Node Reselection Time

To enable a tractable analysis for the reselection time, we adopt a heuristic assumption that when $d \gg s$, the distances $|a_i(t)| \approx p_i(t)$. Therefore, the first hitting time τ in (8) becomes:

$$\tau = \inf\left\{t \geq 0 : \max_{1\leq i\leq L} p_i(t) > \min_{L+1\leq i\leq K} p_i(t)\right\}. \quad (27)$$

In addition, we define the pairwise first hitting time between nodes a_i and a_j, $1 \leq i \leq L$ and $L+1 \leq j \leq K$, as $\tau_{i,j} = \inf\{t \geq 0 : p_i(t) > p_j(t)\}$. The reselection time τ can be rewritten in terms of $\tau_{i,j}$ as $\tau = \min_{i,j}(\tau_{i,j})$, and the distribution of τ is given by $H(t) = \Pr(\tau < t) = 1 - \Pr(\min_{i,j}(\tau_{i,j}) > t)$. Using the Fréchet's inequality [48], $H(t)$ is lower-bounded as:

$$H(t) \geq \max_{i,j} \Pr(\tau_{i,j} < t) = \max_{i,j} H_{i,j}(t) = H_L(t), \quad (28)$$

where $H_{i,j}(t)$ is the first hitting time distribution of nodes a_i and a_j, and for simplicity we denote $H_L(t) \equiv H_{L,L+1}(t)$. The second equality in (28) is clear as $H_L(t)$ is the first hitting time distribution between two nearest nodes from the active and the inactive sets. Assume that at time $t = 0$ the system completes the previous node selection, and that the node motions are in steady state,[9] such that the position of an arbitrary node is uniformly distributed within the cluster area $\mathcal{U}$. Based on these assumptions, an approximation for the first hitting time distribution $H_L(t)$ is given by the following proposition. In particular, the derived approximation is especially accurate at the left tail of the distribution function, i.e., at small $\epsilon = H_L(t_\epsilon)$, which is of the most interest to initiate timely cluster reselection and capture most of the reselection events (27).

[9]In practical system, when the system detects a hitting event defined by (8), there exists operational latency to perform the actual node reselection procedures. When the latency is large, the node motions are assumed to be in steady state.

Proposition 2: The first hitting time distribution $H_L(t) = \Pr(\tau_{L,L+1} < t)$ *is approximated by*

$$H_L(t) \approx e^{\frac{c_1^2}{8c_2}} \sum_{n=0}^{K-1} (-1)^n \mathcal{D}_n \cdot (K-n)_{n+1} \left(\frac{2Dt}{s^2 c_2}\right)^{\frac{n+1}{2}}, \quad (29)$$

where $(a)_m = a(a+1)\cdots(a+m)$ *denotes the Pochhammer symbol,* $\mathcal{D}_n = D_{-n-1}\left(c_1/\sqrt{2c_2}\right)$, $D_{(\cdot)}(\cdot)$ *denotes the parabolic cylinder function [40, eq. (9.240)], and the constants*

$$c_1 = 1.095 \quad \text{and} \quad c_2 = 0.7565. \quad (30)$$

Proof: The proof of Proposition 2 is in Appendix D. ■

Using Proposition 2, the reselection time t_ϵ, defined in (9), can be estimated as $t_\epsilon = H^{-1}(\epsilon) \leq H_L^{-1}(\epsilon)$ by solving the inverse function of $H_L(\cdot)$. In addition, numerical results show that the finite summation (29) can be well approximated by taking the first two summands with $n = 0$ and $n = 1$ for a wide range of system settings. This reduces $H_L(t)$ to a quadratic function in $\sqrt{t}$, and the reselection time t_ϵ can be explicitly solved as:

$$t_\epsilon \approx \frac{s^2 c_2 K^{-1} \mathcal{D}_1^{-2}}{4D(K-1)^2}\left(K\mathcal{D}_0^2 - 2(K-1)\mathcal{D}_1 e^{-\frac{c_1^2}{8c_2}\epsilon} - \mathcal{D}_0\sqrt{K^2\mathcal{D}_0^2 - 4K(K-1)\mathcal{D}_1 e^{-\frac{c_1^2}{8c_2}\epsilon}}\right). \quad (31)$$

As the number of nodes K increases, it is clear from (31) that $t_\epsilon \sim O(1/K^2)$ for a fixed probability ϵ. As the node reselection scheme requires path loss measurements from all K users, the feedback overhead is proportional to K. However, using the well-known opportunistic feedback schemes, such as the one discussed in [49], the amount of feedback scales as $\log K$. Thus, overall, the computational complexity (number of channel measurements per unit time) at the transmit head node scales as $O(K^2 \log K)$. When K is large, the computational complexity may be prohibitively large and a node pre-selection is needed to randomly choose a subset of the K nodes as the candidates in the node selection scheme. In the pre-selection phase, each node is independently selected with a probability K'/K reducing the average number of candidate nodes to K'. The average achievable rate after the pre-selection is given by Proposition 1 with K replaced by K'. The value of K' needs to be configured to trade off between the induced rate degradation and the computational complexity. As shown in the settings of Fig. 2, a pre-selection with $K' = 40$ nodes exploits most of the RD-MIMO performance, and only marginal improvement can be obtained with a larger K'.

Next, we show the impact of the number of transmit nodes on the cluster reconfiguration rate, and we use the reselection time (31) obtained for the Brownian motion model to estimate the reselection time for the random walk mobility model. In the cases of the random walk, the walk step per unit time δ_t is set to $\Delta_a = 0.05$ meters with the instantaneous velocities of $v_{\text{node}} = \Delta_a/\delta_t = 1$, 5, and 10 meters/second (which correspond to 3.6 km/h, 18 km/h, and 36 km/h, respectively). Accordingly, the diffusion coefficient D is set such that the mean square displacements of the Brownian motion and the

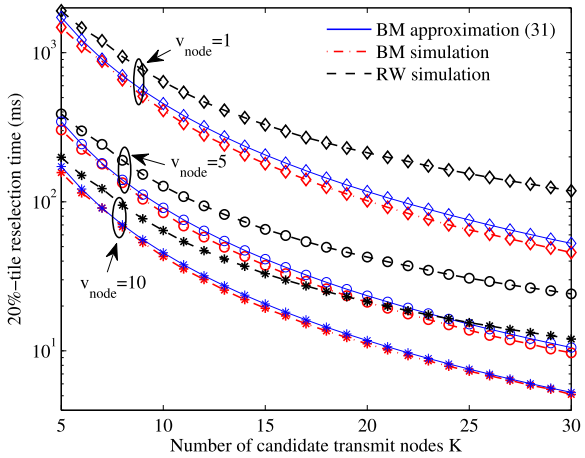


Fig. 4. Comparison of 20%-tile reselection time between Brownian Motion (BM) and Random Walk (RW) models. The walk step of RW is $\Delta_a = 0.05$ meters with various walk speeds $v_{\text{node}} = \Delta_a/\delta_t$. The diffusion coefficient D is set as $D = \Delta_a^2/\delta_t = \Delta_a \cdot v_{\text{node}}$. The number of active transmit nodes is fixed to be $L = 4$.

random walk are equal in each step, i.e., $D\,\delta_t = \Delta_a^2$. Fig. 4 shows the reselection time t_ϵ with $\epsilon = 20\%$ as a function of the number of available nodes K in the cluster area $\mathcal{U}$. The number of active nodes is set to $L = 4$, the inter-cluster distance is $d = 50$ meters, and the edge length of the transmit cluster is $s = 5$ meters. At low velocity of $v_{\text{node}} = 1$, the reselection needs to be performed within 3000 milliseconds when the number of available nodes is small. As K increases, the required reselection time decreases to hundreds of milliseconds. In the case of larger velocities with $v_{\text{node}} = 5$ and 10, the reselection time is below 100 milliseconds for most of the considered values of K. In all cases, the analytical expression (31) provides a useful estimate for the order of magnitude of the reselection time of the corresponding random walk model. Especially, (31) yields a closer estimation when the random walk velocity is larger, which is when the random walk behaves more similar to the Brownian motion. We conclude that in the investigated scenarios, the reselection time is feasible for practical implementation of the RD-MIMO system.

V. Conclusion

As small footprint mobile devices can typically accommodate only a single antenna, and thus have limited ability to cope with severe channel impairments. The proposed reconfigurable distributed MIMO (RD-MIMO) architecture is able to enhance communication between mobile devices, such as is the case in D2D communication scenarios, and between devices and the cellular infrastructure, as is the case in traditional cellular scenarios. With each communicating node temporarily recruiting clusters of adjacent nodes, implementation of temporary distributed MIMO transmission between clusters is enabled as a spatial diversity technique, where the node cooperation is performed via high-speed local wireless links.

Using Random Matrix Theory, the asymptotic average achievable rates of the proposed RD-MIMO architecture were obtained in closed-form expressions. Results show that the

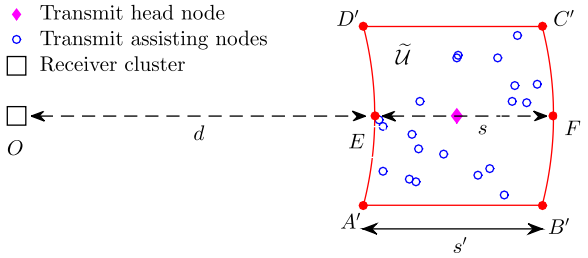


Fig. 5. RD-MIMO with approximate transmit cluster shape.

cluster formed by a selected subset of nodes not only reduces the operational overhead incurred by the cluster cooperation, but may also achieve performance improvement compared to the case of blindly activating all available nodes. In particular, the node selection scheme is an efficient mean to maximize the rate of RD-MIMO in the low SNR regime, where merely 3-6 transmit nodes need to be selected even when there are 10-60 candidate neighboring nodes.

Due to node mobility, the cluster reconfiguration is needed to reselect a new subset of nodes when their channel conditions become more favorable. When the transmit nodes are modeled as Brownian particles, the required time interval to perform node reselection ranges from hundreds of milliseconds at small velocity to dozens of milliseconds at large velocity. The obtained reselection time serves as an estimation for the order of magnitude of the cluster reconfiguration times in other mobility models, such as the random walk model. Importantly, these reconfiguration rates demonstrate that the RD-MIMO technique can be practically implemented in future systems, such as the next generation cellular systems and D2D systems.

Appendix A
Proof of Proposition 1

To enable a tractable analysis for the spectrum of matrix $\Sigma_{(L)}$, instead of dealing with the square area $\mathcal{U}$ shown in Fig. 1, we adopt a further approximate area $\widetilde{\mathcal{U}}$ shown in Fig. 5, which is constructed as follows: keeping the locations of E and F fixed, the edges AD and BC in Fig. 1 are replaced with curves $A'D'$ and $B'C'$, which are parts of the concentric circles with their center at the origin O. The radius of the concentric circles are d and $d+s$, respectively. The vertical coordinates of the edges $A'B'$ and $C'D'$ are $-s/2$ and $s/2$, the same as the corresponding AB and CD, respectively. The length of $A'B'$ and $C'D'$ is now $s' = \sqrt{(d+s)^2 - (s/2)^2} - \sqrt{d^2 - (s/2)^2}$. The approximate cluster $\widetilde{\mathcal{U}}$ effectively projects the transmit nodes onto EF.

Denote $s_{(k)} = d_{(k)} - d$ such that $s_{(1)} \le \cdots \le s_{(K)}$. As the transmit nodes are uniformly distributed within the cluster, $s_{(k)}$ is the k^{th} smallest order statistics of K uniformly distributed random variables over the interval $[0, s]$. It is known that $s_{(k)}$ follows the Beta distribution with density function given by:

$$f_{s_{(k)}}(x) = \frac{x^{k-1}(s-x)^{K-k}}{s^K B(k, K-k+1)}, \quad 0 \le x \le s, \qquad (32)$$

where $B(k, K-k+1) = (k-1)!(K-k)!/K!$ denotes the Beta function. The application of [41, Th. 2.39] requires the knowledge of the Stieltjes transform $\mathcal{G}(\cdot)$ and the Shannon transform $\mathcal{V}(\cdot)$ of the matrix $\Sigma_{(L)}$, which are defined as:

$$\mathcal{G}(\omega) = \mathbb{E}\left[\frac{1}{\lambda_\Sigma - \omega}\right], \qquad \mathcal{V}(\gamma) = \mathbb{E}\left[\log(1+\gamma\ \lambda_\Sigma)\right], \quad (33)$$

where λ_Σ denotes an arbitrary eigenvalue randomly picked from the L eigenvalues of $\Sigma_{(L)}$. Using the density function (32), $\mathcal{G}(\omega)$ is calculated by definition as:

$$\mathcal{G}(\omega) = \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\frac{1}{\sigma_{(l)}-\omega}\right] = \frac{1}{L}\sum_{l=1}^{L}\frac{I_l(\omega)}{B(l, K-l+1)}, \quad (34)$$

where

$$\begin{aligned} I_l(\omega) &= \int_0^1 \frac{z^{l-1}(1-z)^{K-l}(1+z/\theta)^\alpha}{1-w(1+z/\theta)^\alpha}\mathrm{d}z \\ &= \frac{1}{\omega}\int_0^1 \frac{z^{l-1}(1-z)^{K-l}}{1-w(1+z/\theta)^\alpha}\mathrm{d}z - \frac{1}{\omega}\int_0^1 \frac{z^{l-1}}{(1-z)^{l-K}}\mathrm{d}z. \end{aligned} \quad (35)$$

The second integral is identified to be $B(l, K-l+1)$. When α is a positive integer, the integrand of the first integral can be decomposed as a sum of partial fractions [40, eq. (2.102)] as:

$$\frac{z^{l-1}(1-z)^{K-l}(1+z/\theta)^\alpha}{1-w(1+z/\theta)^\alpha} = \frac{1}{\alpha}\sum_{i=1}^{\alpha}\frac{1}{1-\omega_k}\frac{z^{l-1}(1-z)^{K-l}}{1-\frac{\omega_k}{1-\omega_k}\frac{z}{\theta}}, \quad (36)$$

where $\omega_k = |\omega|^{\frac{1}{\alpha}} e^{\frac{2k+1}{\alpha}\pi \mathrm{j}}$, $1 \le k \le \alpha$. Substituting (36) into $I_{l,1}(\omega)$ and applying [40, eq. (3.197.3)], we obtain:

$$\begin{aligned} &\mathcal{G}(\omega) \\ &= \frac{1}{\alpha\,\omega\,L}\sum_{l=1}^{L}\sum_{i=1}^{\alpha}\frac{1}{1-\omega_i}\,{}_2F_1\left(1, l; K+1; \frac{\omega_i}{(1-\omega_i)\theta}\right) - \frac{1}{\omega}. \end{aligned} \quad (37)$$

Interchanging the order of summations and recalling the definition of the generalized hypergeometric functions, we obtain (14).

The Shannon transform of $\Sigma_{(L)}$ is calculated as $\mathcal{V}(\gamma) = \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\log\left(1+\sigma_{(l)}\gamma\right)\right]$. Differentiating $\mathcal{V}(\gamma)$ with respect to γ yields:

$$\frac{d}{d\gamma}\mathcal{V}(\gamma) = \frac{1}{L}\sum_{l=1}^{L}\mathbb{E}\left[\frac{\sigma_{(l)}}{1+\sigma_{(l)}\gamma}\right] = \frac{1}{\gamma} - \frac{1}{\gamma^2}\mathcal{G}\left(-\frac{1}{\gamma}\right), \quad (38)$$

where the second equality is obtained by comparing with (34). Since $\mathcal{V}(0) = 0$, $\mathcal{V}(\gamma)$ in (13) is obtained by direct integration.

Appendix B
Proof of Corollary 1

Assuming $d \to \infty$ while s is fixed, we start the proof of Corollary 1 with a simplified form of $\mathcal{R}_L$ summarized in the following lemma.

Lemma 1: Assume $d \to \infty$ while s fixed. The average achievable rate $\mathcal{R}_L$ is given by:

$$\begin{aligned} \mathcal{R}_L = \zeta\log\left(\frac{\gamma}{\zeta}\mathcal{A}_+(\zeta,\gamma)\right) - \log\mathcal{A}_-(\zeta,\gamma) \\ - \frac{(L+1)\zeta\,\alpha\mathcal{A}_-(\zeta,\gamma)}{2(K+1)\theta\mathcal{A}_+(\zeta,\gamma)} + \mathcal{A}_-(\zeta,\gamma) - 1, \end{aligned} \quad (39)$$

where

$$\begin{aligned} \mathcal{A}_\pm(\zeta,\gamma) = \left(\gamma - \zeta\,\gamma + \sqrt{4\zeta\,\gamma + (\gamma - \zeta - \gamma\,\zeta)^2}\right)/(2\gamma) \\ \pm\,\zeta/(2\gamma). \end{aligned}$$

Proof: With $\omega = -\gamma\ \eta/\zeta$, the fixed-point equation (12) is rewritten as:

$$1 + \frac{1}{\omega}\mathcal{G}\left(\frac{1}{\omega}\right) = \frac{1}{\zeta} + \frac{\omega}{\gamma}. \quad (40)$$

Inserting (14) into the LHS of (40) and applying a series representation [40, eq. (9.14.1)] for ${}_3F_2$, we obtain:

$$\begin{aligned} 1 + \frac{1}{\omega}\mathcal{G}\left(\frac{1}{\omega}\right) &= \frac{1}{\alpha}\sum_{i=1}^{\alpha}\frac{1}{1-|\omega|^{-\frac{1}{\alpha}}e^{\frac{2i+1}{\alpha}\pi\mathrm{j}}} \\ &\quad\times\sum_{k=0}^{\infty}\frac{(L+1)_k}{(K+1)_k(k+1)\theta^k}\frac{1}{(|\omega|^{\frac{1}{\alpha}}e^{-\frac{2i+1}{\alpha}\pi\mathrm{j}}-1)^k} \\ &= \frac{\omega}{\omega-1}\left(1 + \frac{(L+1)\alpha}{2(K+1)\theta}\frac{1}{\omega-1} + O(\theta^{-2})\right), \end{aligned} \quad (41)$$

where the second equality is obtained by interchanging the order of summations. Letting $\theta \to \infty$ and taking the leading order term of (41), the fixed-point equation (40) can be approximated as $\omega/(\omega-1) = 1/\zeta + \omega/\gamma$, which has an explicit solution $\eta = -\omega\,\zeta/\gamma = \mathcal{A}_-(\zeta,\gamma)$.

Next, we derive an approximation for the Shannon transform $\mathcal{V}(\gamma)$ by applying [40, eq. (9.14.1)] to the integrand of (13). Specifically, we have:

$$\begin{aligned} &\frac{1}{\alpha}\sum_{i=1}^{\alpha}\frac{x^{\frac{1}{\alpha}-1}}{x^{\frac{1}{\alpha}} - e^{\frac{2i+1}{\alpha}\pi\mathrm{j}}} \\ &\qquad\times{}_3F_2\left(1, 1, L+1; 2, K+1; \frac{1}{(x^{\frac{1}{\alpha}}e^{-\frac{2i+1}{\alpha}\pi\mathrm{j}}-1)\theta}\right) \\ &= \frac{1}{1+x} - \frac{(L+1)\alpha}{2(K+1)\theta}\frac{1}{(1+x)^2} + O(\theta^{-2}). \end{aligned} \quad (42)$$

Direct integration of (42) yields $\mathcal{V}(\gamma) = \log(1+\gamma) - \frac{(L+1)\alpha}{2(K+1)\theta}\frac{\gamma}{\gamma+1} + O(\theta^{-2})$. Combining $\eta = \mathcal{A}_-(\zeta,\gamma)$, we obtain (39). ∎

Proof of Corollary 1: In the high SNR regime ($\gamma \to \infty$), the asymptotic expansion of $\sqrt{4\zeta/\gamma + (1-\zeta-\zeta/\gamma)^2}$ yields:

$$\begin{aligned} &\sqrt{\frac{4\zeta}{\gamma} + \left(1-\zeta-\frac{\zeta}{\gamma}\right)^2} \\ &\quad= \begin{cases} \zeta - 1 + \dfrac{\zeta(\zeta+1)}{\zeta-1}\dfrac{1}{\gamma} + O(\gamma^{-2}), & \zeta > 1 \\ 1 - \zeta + \dfrac{\zeta(\zeta+1)}{1-\zeta}\dfrac{1}{\gamma} + O(\gamma^{-2}), & \zeta \le 1. \end{cases} \end{aligned} \quad (43)$$

Inserting (43) into $\mathcal{A}_{\pm}(\zeta,\gamma)$ in Lemma 1 and ignoring the infinitesimals as $\gamma\to\infty$, we obtain the leading order terms of $\mathcal{A}_{+}$ and $\mathcal{A}_{-}$ as:

$$\mathcal{A}_{+}(\zeta,\gamma)=\begin{cases}\dfrac{\zeta^2}{(\zeta-1)\gamma}, & \zeta>1\\ 1-\zeta, & \zeta\le 1,\end{cases}$$
$$\mathcal{A}_{-}(\zeta,\gamma)=\begin{cases}\dfrac{\zeta}{(\zeta-1)\gamma}, & \zeta>1\\ 1-\zeta, & \zeta\le 1.\end{cases}\tag{44}$$

Substituting (44) into (39), after algebraic manipulations we obtain (15). The proof for the low SNR regime is similar by noting that the expansion of $\sqrt{4\zeta\,\gamma+(\gamma-\zeta-\gamma\,\zeta)^2}$ at $\gamma=0$ equals:

$$\sqrt{4\zeta\,\gamma+(\gamma-\zeta-\gamma\,\zeta)^2}=\zeta+(1+\zeta)\gamma+o(\gamma^2).\tag{45}$$

The rate approximation (16) follows by inserting (45) into $\mathcal{A}_{\pm}(\zeta,\gamma)$. ∎

Appendix C
Asymptotic Rate With Blind Transmitters

Proof of Corollary 2: When $L=K$, recalling the series representation [40, eq. (9.14.1)] of the generalized hypergeometric function ${}_3F_2$, $\mathcal{G}(\cdot)$ in (14) reduces to:

$$\begin{aligned}\mathcal{G}(\omega)&=-\frac{1}{\omega}+\frac{1}{\alpha\,\omega}\sum_{i=1}^{\alpha}\frac{1}{1-|\omega|^{\frac{1}{\alpha}}e^{\frac{2i+1}{\alpha}\pi\mathrm{j}}}\\&\qquad\times{}_2F_1\left(1,1;2;\frac{1}{(|\omega|^{-\frac{1}{\alpha}}e^{-\frac{2i+1}{\alpha}\pi\mathrm{j}}-1)\theta}\right)\\&=\frac{1}{\omega}\int_0^1\frac{1}{1-\omega(1+t/\theta)^{\alpha}}\mathrm{d}t-\frac{1}{\omega},\end{aligned}\tag{46}$$

where the second equality is due to the integral representation [40, eq. (9.111)] for ${}_2F_1$ and applying the partial fractional decomposition. By variable substitution $z=(1+t/\theta)^{\alpha}$ and applying [40, eq. 3.194.5], we obtain:

$$\begin{aligned}\mathcal{G}(\omega)&=\frac{\theta}{\alpha\,\omega}\int_0^{\left(\frac{\theta+1}{\theta}\right)^{\alpha}}\frac{z^{\frac{1}{\alpha}-1}}{1-\omega\,z}\mathrm{d}z-\frac{\theta}{\alpha\,\omega}\int_0^1\frac{z^{\frac{1}{\alpha}-1}}{1-\omega\,z}\mathrm{d}z\\&=\frac{\theta+1}{\omega}{}_2F_1\left(1,\frac{1}{\alpha};1+\frac{1}{\alpha};\left(\frac{\theta+1}{\theta}\right)^{\alpha}\omega\right)\\&\quad-\frac{\theta}{\omega}{}_2F_1\left(1,\frac{1}{\alpha};1+\frac{1}{\alpha};\omega\right)-\frac{1}{\omega}.\end{aligned}\tag{47}$$

Using the recursive relations [40, eqs. (9.132.2) and (9.137.4)], (47) becomes (24).

Following similar procedures, $\mathcal{V}(\cdot)$ in (23) with $L=K$ can be reduced to:

$$\begin{aligned}\mathcal{V}(\gamma)&=(\theta+1)\log\left(1+\left(\frac{\theta}{\theta+1}\right)^{\alpha}\gamma\right)-\theta\log(1+\gamma)\\&\quad+\frac{\alpha\,\theta\,\gamma}{\alpha-1}\left\{{}_2F_1\left(1,1-\frac{1}{\alpha};2-\frac{1}{\alpha};-\gamma\right)-\left(\frac{\theta}{\theta+1}\right)^{\alpha-1}\right.\\&\qquad\left.\times{}_2F_1\left(1,1-\frac{1}{\alpha};2-\frac{1}{\alpha};-\left(\frac{\theta}{\theta+1}\right)^{\alpha}\gamma\right)\right\}.\end{aligned}\tag{48}$$

Using the recursive relation [40, eq. (9.137.4)], we obtain (23). ∎

Proof of Corollary 3: Letting $\omega(\gamma)=-\gamma\,\eta/\zeta$, the fixed-point equation (12) with $\mathcal{G}(\cdot)$ given by (24) can be rewritten as:

$$\mathcal{F}\big(\omega(\gamma)\big)-\frac{\omega(\gamma)}{\gamma}=\frac{1}{\zeta}-1.\tag{49}$$

According to [41], η is an implicit function of γ and defined as $\eta=\mathbb{E}[1/(1+\gamma\,\lambda_{\Sigma})]$. Therefore, $\omega(\gamma)$ is a decreasing function in γ and we have:

$$\omega(\infty)<\omega(\gamma)\le\omega(0)=0,\quad 0\le\gamma<\infty.\tag{50}$$

As indicated by the inequality (50), ω lies on the negative axis. For $\omega\le 0$, it is easy to verify that $\mathcal{F}(\omega)$ is a decreasing function of ω and $\mathcal{F}(0)=-1$. As $\gamma\to\infty$, $\lim_{\gamma\to\infty}\omega(\gamma)/\gamma=0$, and solving $\mathcal{F}\big(\omega(\infty)\big)=-1+1/\zeta$ with $\zeta>1$ yields $\omega(\infty)$ bounded away from infinity. Since $\mathcal{F}(\omega)$ is a smooth function around the origin, by increasing ζ, $\omega(\infty)$ can be made arbitrarily small. Therefore, when ζ is large, the solution of (49) satisfies $\omega(\gamma)\approx 0$ and $\mathcal{F}(\omega)$ can be approximated by the Taylor expansion at $\omega=0$ such that:

$$\mathcal{F}(\omega)=-1-\frac{\theta}{\alpha-1}\left(1-\left(\frac{\theta}{\theta+1}\right)^{\alpha-1}\right)\omega+\mathcal{O}\left(\omega^2\right).\tag{51}$$

Inserting (51) into (49) and solving for η yields:

$$\eta=\frac{1}{1+\frac{\theta}{\alpha-1}\left(1-\left(\frac{\theta}{\theta+1}\right)^{\alpha-1}\right)\gamma}.\tag{52}$$

Similarly, taking Taylor expansion of $\mathcal{F}(\cdot)$ and of $\log(\cdot)$ at $\omega=0$ in (11), the rate $\mathcal{R}_K$ becomes:

$$\begin{aligned}\mathcal{R}_K&=\log\frac{1}{\eta}+\frac{\theta}{\alpha-1}\left(1-\left(\frac{\theta}{\theta+1}\right)^{\alpha-1}\right)\eta\,\gamma\\&\quad+\eta-1+\mathcal{O}\left(\omega^2\right)=\log\frac{1}{\eta}+\mathcal{O}\left(\omega^2\right),\end{aligned}\tag{53}$$

where the second equality is due to (52). By inserting (52) into (53) and ignoring the infinitesimal, we obtain (26). ∎

Appendix D
Proof of Proposition 2

The conditional survival probability of the process $\tau_{L,L+1}$ given initial condition $p_L(0)=f$ and $p_{L+1}(0)=h$ is defined as: $S(t|f,h)=\Pr\{\tau_{L,L+1}>t|f,h\}$, where $0\le f\le h\le s$. According to [50, eq. (17)], the survival probability $S(t|f,h)$ can be decomposed into $S(t|f,h)=S_1(t|(f+h)/\sqrt{2})\cdot S_1(t|(h-f)/\sqrt{2})$, where $S_1(t|x)$ is given by [51, eq. (X.5.8)] as:

$$\begin{aligned}S_1(t|x)&=\sum_{m=-\infty}^{\infty}\left\{\Phi\left(\frac{(2m+1)R-x}{\sqrt{Dt}}\right)-\Phi\left(\frac{2mR-x}{\sqrt{Dt}}\right)\right.\\&\quad\left.-\Phi\left(\frac{(2m+1)R+x}{\sqrt{Dt}}\right)+\Phi\left(\frac{2mR+x}{\sqrt{Dt}}\right)\right\},\end{aligned}\tag{54}$$

where $R = \sqrt{2}s$, $\Phi(\cdot) = (1 + \mathrm{erf}(x/\sqrt{2}))/2$ denotes the standard Gaussian distribution function, and $\mathrm{erf}(x) = 2/\sqrt{\pi}\int_0^x e^{-t^2}\mathrm{d}t$ is the error function. The expression (54) is known to capture the small time behavior of the first hitting process $\tau_{L,L+1}$, where the infinity summation (54) converges quickly at small t. In particular, we keep the term $m = 0$ in (54) and notice that the Gaussian distribution $\Phi((R \pm x)/\sqrt{Dt}) \approx 1$ when R is significantly larger than the mean displacement of the Brownian motion $\sqrt{Dt}$ at small time. Due to the symmetry of the Gaussian distribution, $S(t|f,h)$ can be approximated as

$$S(t|f,h) \approx \mathrm{erf}\left(-\frac{f+h}{2\sqrt{Dt}}\right)\mathrm{erf}\left(-\frac{h-f}{2\sqrt{Dt}}\right). \tag{55}$$

Using the non-linear least squares method, the error function can be approximated as $\mathrm{erf}(x) = 1-\exp(-c_1x-c_2x^2)$, where the constants c_1 and c_2 in (30) are found in [52]. Therefore, the conditional first hitting probability $H_L(t|f,h) = 1 - S(t|f,h)$ can be approximated as

$$\begin{aligned} H_L(t|f,h) \approx & \exp\left(-c_1\frac{h-f}{2\sqrt{Dt}} - c_2\frac{(h-f)^2}{4Dt}\right) \\ & + \exp\left(-c_1\frac{f+h}{2\sqrt{Dt}} - c_2\frac{(f+h)^2}{4Dt}\right) \\ & - \exp\left(-c_1\frac{h}{\sqrt{Dt}} - c_2\frac{(h-f)^2+(h+f)^2}{4Dt}\right). \end{aligned} \tag{56}$$

Since f and h are realizations of two consecutive ordered statistics, with high probability, we have $f \approx h$. Therefore, at small t, the exponential functions decays quickly at $(f+h)/\sqrt{t}$, while the first term on the RHS of (56) dominates $H_L(t|f,h)$. By ignoring the other terms and integrating (56) over the density [53, eq. (2.3.4)] of $r = h - f$, we obtain

$$H_L(t) = \frac{K}{s^K}\int_0^s (s-r)^{K-1}\exp\left(-\frac{c_1r}{2\sqrt{Dt}} - \frac{c_2r^2}{4Dt}\right)\mathrm{d}r. \tag{57}$$

Since the integrand of (57) decays quickly as $r \to \infty$, we approximate the integral (57) by letting the upper limit of the integral tend to infinity. The proof is completed by using binomial expansion and applying integral identity [40, eq. (3.462.1)].

REFERENCES

[1] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 3, pp. 550–560, Apr. 2002.

[4] D. Gesbert, H. Bölcskei, D. A. Gore, and A. J. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 1926–1934, Dec. 2002.

[5] R. Müller, "A random matrix model of communication via antenna arrays," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2495–2506, Sep. 2002.

[6] M. Chiani, M. Z. Win, and A. Zanella, "On the capacity of spatially correlated MIMO Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2363–2371, Oct. 2003.

[7] A. L. Moustakas, S. H. Simon, and A. M. Sengupta, "MIMO capacity through correlated channels in the presence of correlated interferers and noise: A (not so) large N analysis," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2545–2561, Oct. 2003.

[8] J. Hoydis, R. Couillet, and M. Debbah, "Asymptotic analysis of double-scattering channels," in *Proc. ASILOMAR*, Pacific Grove, CA, USA, Nov. 2011, pp. 1935–1939.

[9] Z. Zheng, L. Wei, R. Speicher, R. Müller, J. Hämäläinen, and J. Corander, "On the finite-SNR diversity-multiplexing tradeoff in large Rayleigh product channels," in *Proc. IEEE Int. Symp. Inf. Theory*, Hong Kong, Jun. 2015, pp. 2593–2597.

[10] Z. Zheng, L. Wei, R. Speicher, R. Müller, J. Hämäläinen, and J. Corander, "Asymptotic analysis of Rayleigh product channels: A free probability approach," *IEEE Trans. Inf. Theory*, to be published. [Online]. Available: http://arxiv.org/abs/1502.05516

[11] D. Aktas, M. N. Bacha, J. S. Evans, and S. V. Hanly, "Scaling results on the sum capacity of cellular networks with MIMO links," *IEEE Trans. Inf. Theory*, vol. 52, no. 7, pp. 3264–3274, Jul. 2006.

[12] W. Roh and A. Paulraj, "MIMO channel capacity for the distributed antenna systems," in *Proc. IEEE Veh. Technol. Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 706–709.

[13] H. Dai, "Distributed versus co-located MIMO systems with correlated fading and shadowing," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4. Toulouse, France, May 2006, pp. 561–564.

[14] J. N. Laneman and G. W. 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.

[15] A. Özgür, O. Lévêque, and D. N. C. Tse, "Spatial degrees of freedom of large distributed MIMO systems and wireless Ad Hoc networks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 202–214, Feb. 2013.

[16] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, Aug. 2010.

[17] C. Zhong, K.-K. Wong, and S. Jin, "Capacity bounds for MIMO Nakagami-m fading channels," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3613–3623, Sep. 2009.

[18] M. Matthaiou, N. D. Chatzidiamantis, and G. K. Karagiannidis, "A new lower bound on the ergodic capacity of distributed MIMO systems," *IEEE Signal Process. Lett.*, vol. 18, no. 4, pp. 227–230, Apr. 2011.

[19] M. Matthaiou, C. Zhong, M. R. McKay, and T. Ratnarajah, "Sum rate analysis of ZF receivers in distributed MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 180–191, Feb. 2013.

[20] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, "Spectral efficiency of distributed MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2112–2127, Oct. 2013.

[21] M. Chiani, M. Z. Win, and H. Shin, "MIMO networks: The effects of interference," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 336–349, Jan. 2010.

[22] L. Dai, "A comparative study on uplink sum capacity with co-located and distributed antennas," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1200–1213, Jun. 2011.

[23] F. Heliot, R. Hoshyar, and R. Tafazolli, "An accurate closed-form approximation of the distributed MIMO outage probability," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 5–11, Jan. 2011.

[24] J. Zhang, C.-K. Wen, S. Jin, X. Gao, and K.-K. Wong, "On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 133–148, Feb. 2013.

[25] J. Gan, Y. Li, S. Zhou, and J. Wang, "On sum rate and power consumption of multi-user distributed antenna system with circular antenna layout," *EURASIP J. Wireless Commun. Netw.*, vol. 2007, p. 089780, Aug. 2007.

[26] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum rate characterization of joint multiple cell-site processing," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4473–4497, Dec. 2007.

[27] O. Simeone, O. Somekh, H. V. Poor, and S. Shamai (Shitz), "Distributed MIMO systems for nomadic applications over a symmetric interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5558–5574, Dec. 2009.

[28] A. Yang, Y. Jing, C. Xing, Z. Fei, and J. Kuang, "Performance analysis and location optimization for massive MIMO systems with circularly distributed antennas," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5659–5671, Oct. 2015.
[29] S.-R. Lee, S.-H. Moon, J.-S. Kim, and I. Lee, "Capacity analysis of distributed antenna systems in a composite fading channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1076–1086, Mar. 2012.
[30] D. Wang, X. You, J. Wang, Y. Wang, and X. Hou, "Spectral efficiency of distributed MIMO cellular systems in a composite fading channel," in *Proc. IEEE Int. Conf. Commun.*, Beijing, China, May 2008, pp. 1259–1264.
[31] G. N. Kamga, M. Xia, and S. Aïssa, "Spectral-efficiency analysis of massive MIMO systems in centralized and distributed schemes," *IEEE Trans. Commun.*, vol. 64, no. 5, pp. 1930–1941, May 2016.
[32] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.
[33] "More than 50 billion connected devices," Ericsson, Stockholm, Sweden, White Paper, Feb. 2011.
[34] H. Yoon and J. Kim, "Collaborative streaming-based media content sharing in WiFi-enabled home networks," *IEEE Trans. Consum. Electron.*, vol. 56, no. 4, pp. 2193–2200, Nov. 2010.
[35] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE J. Sel. Areas Commun.*, vol. 22, no. 6, pp. 1089–1098, Aug. 2004.
[36] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.
[37] *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures (Release 13)*, document TS 36.213 V13.2.0, Tech. Spec. Group Radio Access Network, 2016.
[38] S. Sanayei and A. Nosratinia, "Capacity of MIMO channels with antenna selection," *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 4356–4362, Nov. 2007.
[39] D. Bai, P. Mitran, S. S. Ghassemzadeh, R. R. Miller, and V. Tarokh, "Rate of channel hardening of antenna selection diversity schemes and its implication on scheduling," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4353–4365, Oct. 2009.
[40] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed. New York, NY, USA: Academic, 2007.
[41] A. M. Tulino and S. Verdú, *Random Matrix Theory and Wireless Communications*. Delft, The Netherlands: Now Publishers, 2004.
[42] A. Tulino, A. Lozano, and S. Verdú, "Impact of antenna correlation on the capacity of multiantenna channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 7, pp. 2491–2509, Jul. 2005.
[43] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic MIMO channel model with joint correlation of both link ends," *IEEE Trans. Wireless Commun.*, vol. 5, no. 1, pp. 90–100, Jan. 2006.
[44] W. Hachem, O. Khorunzhiy, P. Loubaton, J. Najim, and L. Pastur, "A new approach for mutual information analysis of large dimensional multi-antenna channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 9, pp. 3987–4004, Sep. 2008.
[45] V. Raghavan and A. M. Sayeed, "Sublinear capacity scaling laws for sparse MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 345–364, Jan. 2011.
[46] V. Raghavan, A. M. Sayeed, and V. V. Veeravalli, "Semiunitary precoding for spatially correlated MIMO channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1284–1298, Mar. 2011.
[47] Ö. Oyman, R. U. Nabar, H. Bölcskei, and A. J. Paulraj, "Characterizing the statistical properties of mutual information in MIMO channels," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2784–2795, Nov. 2003.
[48] T. Hailperin, *Boole's Logic and Probability: A Critical Exposititon From the Standpoint of Contemporary Algebra, Logic, and Probability Theory*, 2nd ed. Amsterdam, The Netherlands: North Holland, 1986.
[49] R. Agarwal, V. R. Majjigi, Z. Han, R. Vannithamby, and J. M. Cioffi, "Low complexity resource allocation with opportunistic feedback over downlink OFDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1462–1472, Oct. 2008.
[50] R. Groenevelt, E. Altman, and P. Nain, "Relaying in mobile ad hoc networks: The Brownian motion mobility model," *Wireless Netw.*, vol. 12, no. 5, pp. 561–571, Sep. 2006.
[51] W. Feller, *An Introduction to Probability Theory and Its Applications*, 2nd ed. New York, NY, USA: Wiley, 1971.
[52] W.-J. Tsay, C. J. Huang, T.-T. Fu, and I.-L. Ho, "Maximum likelihood estimation of censored stochastic frontier models: An application to the three-stage DEA method," Inst. Econ., Academia Sinica, Taipei, Tech. Rep. 09-A003, 2009.
[53] H. A. David, *Order Statistics*, 2nd ed. New York, NY, USA: Wiley, 1981.

Zhong Zheng (S'10–M'15) received the B.Eng. degree from the Beijing University of Technology, China, in 2007, the M.Sc. degree from the Helsinki University of Technology, Finland, in 2010, and the D.Sc. degree from Aalto University, Finland, in 2015. He is currently a Post-Doctoral Researcher with the Department of Computer Science, University of Texas at Dallas, USA. His research interests include MIMO communications, secure communications, heterogeneous networks, random matrix theory, and free probability theory.

Zygmunt J. Haas (S'84–M'88–SM'90–F'07) received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1988. In 1988, he joined AT&T Bell Laboratories in the Network Research Area, where he pursued research in wireless communications, mobility management, fast protocols, optical networks, and optical switching. In 1995, he joined the Faculty with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, USA, where he is currently a Professor Emeritus. He is also a Professor and a Distinguished Chair with the Computer Science Department, The University of Texas at Dallas, Richardson, TX, USA, since 2013. He heads the Wireless Network Laboratory, a research group with extensive contributions and international recognition in the area of ad hoc networks and sensor networks. He has authored over 200 technical conference and journal papers. He holds 18 patents in the areas of wireless networks and wireless communications, optical switching and optical networks, and high-speed networking protocols. His research interests include protocols for mobile and wireless communication and networks, secure communications, and the modeling and performance evaluation of large and complex systems. He was a recipient of a number of awards and distinctions, including best paper awards, the 2012 IEEE ComSoc WTC Recognition Award for outstanding achievements and contribution in the area of wireless communications systems and networks, and the 2016 IEEE ComSoc AHSN Recognition Award for outstanding contributions to securing ad hoc and sensor networks. He has organized several workshops, delivered numerous tutorials at major IEEE and ACM conferences, and served as an Editor of several journals and magazines, including the IEEE/ACM TRANSACTIONS ON NETWORKING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the *IEEE Communications Magazine*, and the *Wireless Networks* (Springer). He has been a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS issues and served as the Chair of the IEEE Technical Committee on Personal Communications.