

Relay Selection in NOMA-Based Diamond Relaying Networks

Noha H. Khattab¹, Samy S. Soliman^{2,3}, M. Saeed Darweesh¹, and Amr A. El-Sherif^{4,5}

¹Wireless Intelligent Networks Center (WINC), Nile University, Giza, Egypt

²Electronics and Electrical Communications Engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt

³University of Science and Technology, Zewail City of Science, Technology and Innovation, Giza, Egypt

⁴School of Engineering, New Giza University, Giza, Egypt

⁵Department of Electrical Engineering, Alexandria University, Alexandria, Egypt

n.khattab@nu.edu.eg, samy.soliman@cu.edu.eg, mdarweesh@nu.edu.eg, amr.elsherif@ngu.edu.eg

Abstract—Diamond relaying networks based on non-orthogonal multiple access (NOMA) can offer higher achievable data rates than the conventional relaying networks. In diamond relaying networks, two relays cooperate to deliver two messages to the destination within only two time slots. This paper investigates the impact of relay selection, from two relay pools, on the performance of such networks. The system performance is analyzed through closed-form and asymptotic expressions of the sum ergodic rate. The paper also presents practical design guidelines of optimum location-based relay selection for IoT-based networks with randomly deployed relays.

Index Terms—Cooperative relaying, ergodic capacity, NOMA, relay selection, sum rate.

I. INTRODUCTION

The Internet of Things (IoT) and its next generations are predicted to be a pivotal use case for fifth-generation (5G) communication networks and beyond [1]. The number of connected devices within a network has been increasing exponentially and is expected to reach 80 billion devices by 2030 [2]. This rapid growth has posed significant challenges to wireless communication networks in fulfilling high connectivity requirements. Non-orthogonal multiple access (NOMA) is considered a promising candidate for future wireless networks due to its potentials to support massive connectivity in IoT applications [3]. Compared to the conventional orthogonal multiple access (OMA) techniques, the advantage of NOMA systems is that multiple receivers, with distinct channel gains, can share the same resource block, such as time, frequency, or code. NOMA is accomplished using superposition coding at the transmitter with different power coefficients assigned to different data symbols. At the receiver, successive interference cancellation (SIC) is used to extract the desired information by eliminating unwanted information of other users according to the allocated power [4], [5]. Several studies have shown that NOMA can support massive connectivity while offering higher spectral efficiency and user fairness when compared to OMA [6], [7].

To further improve cell coverage, spatial diversity, and overall system performance, cooperative relaying has been integrated into NOMA networks [8]. In cooperative NOMA

networks, the source transmits data to the unreachable destination with the help of a decode-and-forward relay; either this relay is a direct reachable user or a dedicated node for relaying [9]. Then, in order to increase the diversity gain, multi-relay networks with relay selection (RS) strategies have been studied in [10] - [13]. In particular, the authors in [11] proposed a two-stage RS method that was tested under various quality-of-service (QoS) requirements and outperforms the scheme proposed in their previous work [10]. In IoT systems, cooperation plays an essential role in the communication between the sources and the destinations, with several relaying nodes in between. In [13], the authors studied RS in NOMA-based IoT systems. Closed-form expressions of the outage probabilities were derived in the scenario where a base station sends messages to two sensors with an eavesdropper. The best-relay selection method was recommended to enhance the system performance.

Unlike conventional cooperative relaying network models, which only send a single symbol to the destination during two-time intervals, authors in [14] proposed a new network model known as network diamond relaying (NDR). NOMA-based NDR network model could simultaneously convey two symbols to the destination within only two time transmission phases, thus increasing the multiplexing gain. Following the downlink NOMA principle, the source transmits two superimposed symbols with different power allocations to two intermediate relays with varying distances from the source. Then, the relays forward the decoded signals to the destination using NOMA in an uplink mode. As a result, this proposed system can reap the benefits of NOMA in both downlink and uplink transmission.

Authors in [14] showed that the NOMA-based NDR network outperforms the other conventional cooperative schemes that rely on the maximal ratio combining (MRC) and the max-min criterion in terms of the sum rate. Then, they extended their work in [15], in which a closed-form expression of the ergodic sum rate has been derived. Also, power allocation algorithms have been introduced under both the statistical and instantaneous channel state information assumptions. In [16], a joint power allocation approach for improving the achievable

rate of NOMA-based diamond relaying networks has been introduced. Simulation results have shown that the proposed joint power allocation method can offer a higher achievable rate in the NDR networks than that obtained in [15]. Finally, an efficient algorithm has been presented in [17] to offer a higher achievable rate than other state-of-the-art approaches.

Most of the previous work on NOMA-based NDR networks only focused on investigating the system's performance and was limited to a single-relay scenario. They have not examined the benefits of multiple relays with the advantage of RS to improve transmission reliability, limiting RS's functionality in practical systems. Besides, a study case for practical design guidelines considering the effect of the location of the selected relays should be investigated.

In this work, motivated by the observations of the literature, the NOMA-based NDR network model has been extended into the IoT network deployment scenario. The modified model in this paper considers randomly distributed relaying nodes between the source and the destination and studies the effect of RS on the overall sum rate. A proper RS scheme is used for selecting two relays at optimum locations to assess the end-to-end transmission process. Also, the impact of increasing the number of relays at either optimum or non-optimum relay locations is investigated. A practical design guideline of the optimum location-based relay selection for IoT-based networks is investigated. Finally, the analytical expressions for the sum rate of the considered system are derived under Rayleigh fading channels.

The remainder of this paper is organized as follows. The system model, with relay selection, is introduced in Section II. Analytical expressions of the ergodic sum rate are presented in Section III. Numerical results are shown in Section IV, and Section V concludes the paper.

II. SYSTEM MODEL

This paper considers a network with a single base station source node (S), a single destination node (D), and two sets of relays in between. Each set of relays exist at a different distance from S and D . This forms two communication branches between the source and the destinations, referred to as the upper and lower branches, as shown in Fig. 1. It is assumed that all nodes are equipped with a single antenna hence operate in a half-duplex mode. One relay from each set of relays is chosen to cooperate with the transmission of data symbols from the source to the destination. Assume that R_1 is the best relay selected from the upper branch relays $R_{1,i}$, where $1 \leq i \leq K$, and R_2 is the relay selected from the lower branch relays $R_{2,j}$, where $1 \leq j \leq M$. Accordingly, a pair of symbols is transmitted from the source to the destination over two phases, which may be considered a downlink NOMA phase and an uplink NOMA phase.

The channel coefficients for the links between the source and R_1 , the source and R_2 , R_1 and the destination as well as R_2 to the destination are denoted, respectively, by h_{SR_1} , h_{SR_2} , h_{R_1D} and h_{R_2D} . The channel gains h_i , where $i \in \{SR_1, SR_2, R_1D, R_2D\}$ are assumed to be Rayleigh faded,

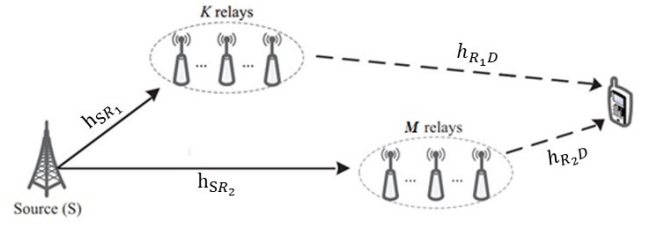


Fig. 1. Network diamond relaying (NDR) system model

i.e. $h_i \sim \mathcal{CN}(0, \lambda_i)$, with zero mean and variance $\lambda_i = \eta d_i^{-n}$ where η represents the propagation constant, n is the path loss exponent and d_i is the distance between the two terminals of the channel. Accordingly, the random variables $|h_i|^2$ follow exponential distributions with means λ_i .

The following section will introduce the mathematical model and expressions that relate the achievable data rate to the system parameters based on the specified system.

A. The Downlink NOMA Phase

In the downlink NOMA phase, the source transmits two superimposed data symbols, s_1 and s_2 . The transmitted signal representing the superimposed signals, s , is expressed as $s = \sqrt{\alpha_1 P_t} s_1 + \sqrt{\alpha_2 P_t} s_2$, where α_1 and α_2 are the power allocation coefficients and P_t is the total transmission power from the base station source. The first symbol, s_1 , is to be decoded at R_1 , the relay selected from the upper branch pool of relays. Similarly, the second symbol s_2 is to be decoded at R_2 . Both decoded symbols will be transmitted simultaneously, each from one of the two relays, to the destination during the uplink NOMA phase.

The distance between the source, S , and R_1 is assumed to be shorter than the distance between the source and R_2 . By assuming $d_{SR_1} < d_{SR_2}$, it is possible to consider that the average channel gains are such that $\lambda_{SR_1} > \lambda_{SR_2}$. Accordingly, $\mathbb{E}\{|h_{SR_1}|^2\} > \mathbb{E}\{|h_{SR_2}|^2\}$.

The received signal at the selected upper branch relay, R_1 , can be given as

$$\begin{aligned} y_1 &= h_{SR_1} s + w_1 \\ &= h_{SR_1} (\sqrt{\alpha_1 P_t} s_1 + \sqrt{\alpha_2 P_t} s_2) + w_1 \\ &= \underbrace{h_{SR_1} \sqrt{\alpha_1 P_t} s_1}_{\text{desired \& low power}} + \underbrace{h_{SR_1} \sqrt{\alpha_2 P_t} s_2}_{\text{interference \& dominating}} + \underbrace{w_1}_{\text{noise}} \end{aligned} \quad (1)$$

where w_1 represents zero mean AWGN, with average noise power σ^2 , at R_1 . It is assumed that the first symbol is to be extracted for forwarding at relay R_1 , while the second symbol at relay R_2 . It is also assumed that the first symbol, s_1 , is extracted at R_1 through applying SIC by decoding the s_2 first, while R_2 decodes for s_2 directly. Hence, the power allocation coefficients must be such that $\alpha_2 > \alpha_1$. The signal-to-noise ratio at R_1 for decoding s_2 , denoted by $\gamma_{s_2 \rightarrow R_1}$, can be expressed as

$$\gamma_{s_2 \rightarrow R_1} = \frac{\alpha_2 P_t |h_{SR_1}|^2}{\alpha_1 P_t |h_{SR_1}|^2 + \sigma^2} = \frac{\alpha_2 \rho_t |h_{SR_1}|^2}{\alpha_1 \rho_t |h_{SR_1}|^2 + 1}, \quad (2)$$

where $\rho_t = \frac{P_t}{\sigma^2}$ represents the ratio between the total transmit signal power to the average noise power at R_1 .

In order to consider s_2 successfully decoded, the data rate for s_2 , denoted as $C_{S_{R_1}}^{s_2}$, on the S -to- R_1 link must be greater than the target data rate, R_{s_2} , for s_2 . This condition can be expressed as

$$\begin{aligned} C_{S_{R_1}}^{s_2} &= \log_2(1 + \gamma_{s_2 \rightarrow R_1}) \\ &= \log_2\left(1 + \frac{\alpha_2 \rho_t |h_{SR_1}|^2}{\alpha_1 \rho_t |h_{SR_1}|^2 + 1}\right) \geq R_{s_2}. \end{aligned} \quad (3)$$

With this condition fulfilled, SIC is completed to decode s_1 . As a result, the signal to noise ratio at R_1 for decoding s_1 , denoted by $\gamma_{s_1 \rightarrow R_1}$, can be written as

$$\begin{aligned} \gamma_{s_1 \rightarrow R_1} &= \frac{\alpha_1 P_t |h_{SR_1}|^2}{\sigma^2} \\ &= \alpha_1 \rho_t |h_{SR_1}|^2. \end{aligned} \quad (4)$$

Thus, the corresponding data rate achieved for symbol s_1 , denoted as $C_{S_{R_1}}^{s_1}$, on the S -to- R_1 link can be expressed as:

$$C_{S_{R_1}}^{s_1} = \log_2(1 + \gamma_{s_1 \rightarrow R_1}) = \log_2(1 + \alpha_1 \rho_t |h_{SR_1}|^2). \quad (5)$$

Simultaneously, R_2 shall acquire the second symbol s_2 , which could be directly decoded, and s_1 could be treated as noise. The received signal at R_2 can be given as

$$\begin{aligned} y_2 &= h_{SR_2} s + w_2 \\ &= h_{SR_2} (\sqrt{\alpha_1 P_t} s_1 + \sqrt{\alpha_2 P_t} s_2) + w_2 \\ &= \underbrace{h_{SR_2} \sqrt{\alpha_1 P_t} s_1}_{\text{interference \& low power}} + \underbrace{h_{SR_2} \sqrt{\alpha_2 P_t} s_2}_{\text{desired \& dominating}} + \underbrace{w_2}_{\text{noise}} \end{aligned} \quad (6)$$

where w_2 represents zero mean AWGN, with average noise power σ^2 , at the relay R_2 . The signal to noise ratio, $\gamma_{s_2 \rightarrow R_2}$, at R_2 for decoding s_2 can be then expressed as

$$\gamma_{s_2 \rightarrow R_2} = \frac{\alpha_2 \rho_t |h_{SR_2}|^2}{\alpha_1 \rho_t |h_{SR_2}|^2 + 1}. \quad (7)$$

The corresponding data rate achieved for symbol s_2 on the S -to- R_2 link, and denoted $C_{S_{R_2}}^{s_2}$, can be given as

$$\begin{aligned} C_{S_{R_2}}^{s_2} &= \log_2(1 + \gamma_{s_2 \rightarrow R_2}) \\ &= \log_2\left(1 + \frac{\alpha_2 \rho_t |h_{SR_2}|^2}{\alpha_1 \rho_t |h_{SR_2}|^2 + 1}\right). \end{aligned} \quad (8)$$

B. The Uplink NOMA Phase

In the second transmission phase, each of the two relays R_1 and R_2 , forwards the decoded symbol, s_1 and s_2 , respectively, to the destination node. Following the uplink NOMA principle, R_1 transmits the data symbol s_1 with transmit power $\alpha_3 P_t$, while R_2 transmits the data symbol s_2 with transmit power $\alpha_4 P_t$. Accordingly, the received signal at the destination is expressed as

$$y_d = \underbrace{\sqrt{\alpha_3 P_t} s_1 h_{R_1 D}}_{\text{From } R_1} + \underbrace{\sqrt{\alpha_4 P_t} s_2 h_{R_2 D}}_{\text{From } R_2 \text{ (dominating)}} + w_3 \quad (9)$$

where w_3 represents zero mean AWGN, with average noise power σ^2 , at the destination D . Based on the described model, the relay R_2 is closer to the destination than the relay R_1 , implying a higher channel gain for the R_2 -to- D link than that for the R_1 -to- D link, i.e. $|h_{R_2 D}|^2 > |h_{R_1 D}|^2$. With appropriate power allocation of $\alpha_3 P_t$ and $\alpha_4 P_t$ such that $\alpha_4 > \alpha_3$, the signal received at the destination is dominated by the signal sent from R_2 . This validated the NOMA considerations for the uplink power domain during this phase of transmission. It is worth noting that the total power allocated for the transmission of both symbols s_1 and s_2 is considered to be P_t , which includes the power utilized at the source and both forwarding relays. Accordingly, it is assumed that $\sum_{i=1}^4 \alpha_i = 1$. It is assumed that the base station performs power allocation to determine the power allocation coefficients, α_i , $i = 1, 2, 3, 4$. It should be noted that optimum power allocation has been an essential aspect of NOMA-based communication models. However, the details of the power allocation algorithms followed at the base station are considered in other work.

Similar to the analysis presented for the downlink NOMA phase, the received signal expressed in (9) is processed to directly decode s_2 , the stronger component, treating s_1 as noise. Therefore, the signal to noise ratio, $\gamma_{s_2 \rightarrow D}$, at the destination for decoding s_2 can be expressed as

$$\gamma_{s_2 \rightarrow D} = \frac{\alpha_4 \rho_t |h_{R_2 D}|^2}{\alpha_3 \rho_t |h_{R_1 D}|^2 + 1}. \quad (10)$$

The achievable data rate, $C_{R_2 D}^{s_2}$, for symbol s_2 on the R_2 -to- D link can be written as

$$C_{R_2 D}^{s_2} = \log_2\left(1 + \frac{\alpha_4 \rho_t |h_{R_2 D}|^2}{\alpha_3 \rho_t |h_{R_1 D}|^2 + 1}\right). \quad (11)$$

Similar to the consideration in (3), successful decoding of s_2 is considered if the achievable data rate is greater than a determined target data rate R_{s_2} , resulting in the condition

$$C_{R_2 D}^{s_2} \geq R_{s_2}, \quad (12)$$

With s_2 decoded, SIC is continued to decode s_1 . As a result, the signal to noise ratio at destination for decoding s_1 , $\gamma_{s_1 \rightarrow D}$, is expressed as

$$\gamma_{s_1 \rightarrow D} = \alpha_3 \rho_t |h_{R_1 D}|^2. \quad (13)$$

The resulting achievable data rate, $C_{R_1 D}^{s_1}$, for symbol s_1 on the R_1 -to- D link can be then expressed as

$$\begin{aligned} C_{R_1 D}^{s_1} &= \log_2(1 + \gamma_{s_1 \rightarrow D}) \\ &= \log_2(1 + \alpha_3 \rho_t |h_{R_1 D}|^2). \end{aligned} \quad (14)$$

To determine the overall system achievable data rate, the achievable rates of both delivered symbols, C_{s_1} and C_{s_2} , are added. According to the authors in [18], the end-to-end achievable data rate of conventional cooperative relaying system is limited by the worst link of the source-relay and relay-destination links. From (5) and (14), it can be obviously observed that the achievable rate of s_1 will be limited by the

minimum of $C_{S_{R_1}}^{s_1}$ and $C_{R_{1D}}^{s_1}$. Similarly, from (3), (8) and (11), the achievable rate of s_2 will be limited to the minimum of $C_{S_{R_1}}^{s_2}$, $C_{S_{R_2}}^{s_2}$ and $C_{R_{2D}}^{s_2}$. Thus, the overall achievable data rate is written as

$$C = \underbrace{\frac{1}{2} \min(C_{S_{R_1}}^{s_1}, C_{R_{1D}}^{s_1})}_{C_{s_1}} + \underbrace{\frac{1}{2} \min(C_{S_{R_1}}^{s_2}, C_{S_{R_2}}^{s_2}, C_{R_{2D}}^{s_2})}_{C_{s_2}}, \quad (15)$$

where the factor $\frac{1}{2}$ in (15) accounts for the fact that each symbol requires two phases to reach the destination.

C. Relay Selection Strategy

As proposed in the system model, relay selection can be employed to enhance the performance of the NDR model. The proposed model adopts a direct strategy in which the highest end-to-end sum achievable data rate in (15) is sought. Accordingly, the k^{th} relay is selected from the upper branch relay pool, and the m^{th} relay is selected from the lower branch relay pool if their combined data rate achieves the maximum sum data rate. This can be expressed as

$$\{k^*, m^*\} = \arg \max_{k \in [1, K], r \in [1, M]} \{C_{k, m}\}. \quad (16)$$

The optimally selected relays are denoted R_1 and R_2 in the analysis presented in (1) - (14). The effect of the numbers of relays, K and M , and the effects of their relative positions to each other as well as relative to the source base station and the destination are studied in the simulation and results section.

In the following section, the ergodic capacity, based on the described model, is defined. Closed-form and asymptotic expressions for the ergodic capacities of both symbols s_1 and s_2 , are also derived.

III. ERGODIC CAPACITY ANALYSIS

In this section, the system performance in terms of the sum ergodic rate is discussed. For the system model discussed in the previous section, the sum ergodic data rate, $\mathbb{E}\{C\}$, i.e. the average sum of the achievable data rate of symbols s_1 and s_2 in (15) can be expressed as

$$\mathbb{E}\{C\} = \mathbb{E}\{C_{s_1}\} + \mathbb{E}\{C_{s_2}\}. \quad (17)$$

The ergodic data rate is calculated by averaging the instantaneous capacity achieved through transmission over a varying fading channel. Since the data rate is dependent on the signal to noise ratio, the ergodic data rate is expressed as

$$\mathbb{E}\{C_x\} = \int_0^\infty \log_2(1+x) f_X(x) dx \quad (18)$$

$$= \frac{1}{\ln 2} \int_0^\infty \frac{1 - F_X(x)}{1+x} dx, \quad (19)$$

where $f_X(x)$ and $F_X(x)$ are respectively the probability density function (PDF) and the cumulative distribution function (CDF) of random random variable X . In the following

subsections, the ergodic data rate for each of the two symbols s_1 and s_2 is calculated.

A. Ergodic data rate of symbol s_1

The instantaneous data rate achieved for symbol s_1 at the receiver can be obtained by substituting (5) and (14) into the first term of (15), resulting in

$$C_{s_1} = \frac{1}{2} \log_2(1 + \min\{\gamma_{s_1 \rightarrow R_1}, \gamma_{s_1 \rightarrow D}\}). \quad (20)$$

Consequently, the ergodic capacity for s_1 can be evaluated through

$$\mathbb{E}\{C_{s_1}\} = \frac{1}{2} \frac{1}{\ln 2} \int_0^\infty \frac{1 - F_U(u)}{1+u} du. \quad (21)$$

where $U = \min\{\gamma_{s_1 \rightarrow R_1}, \gamma_{s_1 \rightarrow D}\}$. Therefore, using (21) to calculate the ergodic data rate of symbol s_1 requires the calculation of the CDF of U . Using order statistics, the CDF of the random variable U can be obtained as

$$\begin{aligned} F_U(u) &= P(\min\{\gamma_{s_1 \rightarrow R_1}, \gamma_{s_1 \rightarrow D}\} < u) \\ &= 1 - P(\gamma_{s_1 \rightarrow R_1} \geq u) P(\gamma_{s_1 \rightarrow D} \geq u) \\ &= 1 - F_{\gamma_{s_1 \rightarrow R_1}}(u) F_{\gamma_{s_1 \rightarrow D}}(u). \end{aligned} \quad (22)$$

Since the random variables $\gamma_{s_1 \rightarrow R_1}$ and $\gamma_{s_1 \rightarrow D}$, defined in (4) and (13), follow exponential distribution with means $\alpha_1 \rho_t \lambda_{SR_1}$ and $\alpha_3 \rho_t \lambda_{R_1 D}$, hence, their CDFs can be, respectively, expressed as

$$F_{\gamma_{s_1 \rightarrow R_1}}(u) = 1 - e^{-\frac{u}{\alpha_1 \rho_t \lambda_{SR_1}}} \quad u > 0. \quad (23)$$

$$F_{\gamma_{s_1 \rightarrow D}}(u) = 1 - e^{-\frac{u}{\alpha_3 \rho_t \lambda_{R_1 D}}} \quad u > 0. \quad (24)$$

Therefore, substituting (23) and (24) into (22), the CDF of random variable U can be given by

$$F_U(u) = 1 - e^{-\frac{u}{\rho_t} (\frac{1}{\alpha_1 \lambda_{SR_1}} + \frac{1}{\alpha_3 \lambda_{R_1 D}})} \quad u > 0. \quad (25)$$

Substituting into (21), the ergodic capacity of s_1 is then obtained as

$$\mathbb{E}\{C_{s_1}\} = \frac{1}{2 \ln 2} \int_0^\infty \frac{e^{-\frac{u}{\rho_t} (\frac{1}{\alpha_1 \lambda_{SR_1}} + \frac{1}{\alpha_3 \lambda_{R_1 D}})}}{1+u} du. \quad (26)$$

The integration in (26) can be solved using [19, eq. (3.352.4)], resulting in

$$\begin{aligned} \mathbb{E}\{C_{s_1}\} &= -\frac{1}{2 \ln 2} e^{-\frac{1}{\rho_t} (\frac{1}{\alpha_1 \lambda_{SR_1}} + \frac{1}{\alpha_3 \lambda_{R_1 D}})} \\ &\quad \times Ei \left(\frac{-1}{\rho_t} \left(\frac{1}{\alpha_1 \lambda_{SR_1}} + \frac{1}{\alpha_3 \lambda_{R_1 D}} \right) \right). \end{aligned} \quad (27)$$

where $Ei(\cdot)$ denotes exponential integral function [19, eq. (8.211.1)].

B. Ergodic data rate of symbol s_2

Similar to the aforementioned analysis, by substituting (3), (8) and (11) into the second term of (15), the instantaneously achievable data rate of symbol s_2 can be expressed as

$$C_{s_2} = \frac{1}{2} \log_2 (1 + \min \{\gamma_{s_2 \rightarrow R_1}, \gamma_{s_2 \rightarrow R_2}, \gamma_{s_2 \rightarrow D}\}). \quad (28)$$

Similar to (21), (28) can be solved by letting $V = \min \{\gamma_{s_2 \rightarrow R_1}, \gamma_{s_2 \rightarrow R_2}, \gamma_{s_2 \rightarrow D}\}$ and considering order statistics, the CDF of random variable V is given by:

$$F_V(v) = 1 - F_{\gamma_{s_2 \rightarrow R_1}}(v)F_{\gamma_{s_2 \rightarrow R_2}}(v)F_{\gamma_{s_2 \rightarrow D}}(v). \quad (29)$$

The CDFs of $\gamma_{s_2 \rightarrow R_1}$, $\gamma_{s_2 \rightarrow R_2}$ and $\gamma_{s_2 \rightarrow D}$ can be evaluated, respectively, as

$$F_{\gamma_{s_2 \rightarrow R_1}}(v) = 1 - e^{-\frac{v}{\alpha_2 - v\alpha_1} \frac{1}{\rho_t \lambda_{SR1}}} \quad 0 < v < \frac{\alpha_2}{\alpha_1}, \quad (30)$$

$$F_{\gamma_{s_2 \rightarrow R_2}}(v) = 1 - e^{-\frac{v}{\alpha_2 - v\alpha_1} \frac{1}{\rho_t \lambda_{SR2}}} \quad 0 < v < \frac{\alpha_2}{\alpha_1} \quad (31)$$

and

$$F_{\gamma_{s_2 \rightarrow D}}(v) = 1 - \frac{\alpha_4 \rho_t \lambda_{R2D}}{\alpha_4 \rho_t \lambda_{R2D} + \alpha_3 \rho_t \lambda_{R1D}} \times e^{-\frac{v}{\alpha_4 \rho_t \lambda_{R2D}}} \quad v > 0. \quad (32)$$

Therefore, the CDF of random variable V can be obtained by substituting (30), (31) and (32) into (29) resulting in

$$F_V(v) = 1 - \frac{\alpha_4 \rho_t \lambda_{R2D}}{\alpha_4 \rho_t \lambda_{R2D} + v\alpha_3 \rho_t \lambda_{R1D}} \times e^{-\frac{v}{\alpha_4 \rho_t \lambda_{R2D}}} \times e^{-\frac{v}{\alpha_2 - v\alpha_1} \left(\frac{1}{\rho_t \lambda_{SR1}} + \frac{1}{\rho_t \lambda_{SR2}} \right)} \quad 0 < v < \frac{\alpha_2}{\alpha_1}. \quad (33)$$

The evaluation of the integral resulting from substituting (33) into (19) is analytically complex. Instead, the ergodic capacity of symbol s_2 is evaluated asymptotically for high SNR values. When $\rho_t \rightarrow \infty$, the CDF of V in (33) can be approximated to

$$F_{V_{\rho_t \rightarrow \infty}}(v) = 1 - \frac{\alpha_4 \lambda_{R2D}}{\alpha_4 \lambda_{R2D} + v\alpha_3 \lambda_{R1D}} \quad v < \frac{\alpha_2}{\alpha_1}. \quad (34)$$

Thus, the PDF of V can be approximated by

$$f_{V_{\rho_t \rightarrow \infty}}(v) = \frac{\alpha_4 \lambda_{R2D} \alpha_3 \lambda_{R1D}}{(\alpha_4 \lambda_{R2D} + v\alpha_3 \lambda_{R1D})^2} \quad v < \frac{\alpha_2}{\alpha_1}. \quad (35)$$

Substituting this PDF expression into (18), the asymptotic ergodic rate of symbol s_2 can be represented as

$$\mathbb{E}\{C_{s_2}\} = \frac{1}{2 \ln 2} \frac{\alpha_4 \lambda_{R2D}}{\alpha_4 \lambda_{R2D} - \alpha_3 \lambda_{R1D}} \times \ln \left(\frac{\alpha_4 \lambda_{R2D} + \frac{\alpha_2}{\alpha_1} \alpha_4 \lambda_{R2D}}{\alpha_4 \lambda_{R2D} + \frac{\alpha_2}{\alpha_1} \alpha_3 \lambda_{R1D}} \right). \quad (36)$$

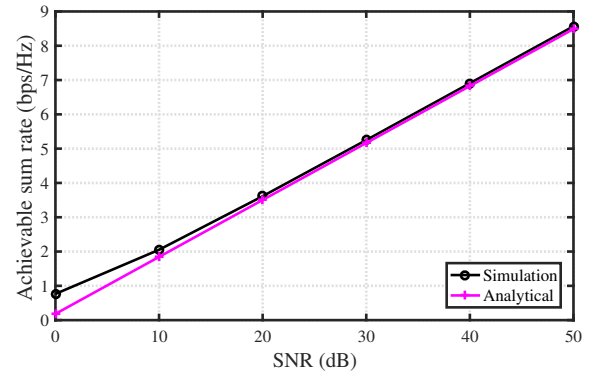


Fig. 2. Achievable ergodic sum rate of NDR system model. Solid markers shown analytical results and open markers show simulation results.

IV. SIMULATION AND RESULTS

In this section, simulation results are presented to demonstrate the performance of the system explained and analyzed in the previous sections. First, Monte Carlo simulations are used to verify the analytical expressions of the ergodic capacity obtained through substituting (27) and (36) into (17). Then, numerical simulation and analytical expressions are used to study the effects of the distances between the source node, the relays, and the destination node on the achievable data rate. Simulations are also used to study the effects of the number of relays in the relay selection pools on the system's ergodic capacity.

In the following simulation examples, a fixed power allocation strategy is adopted, in which $\alpha_1 = 2.5\%$, $\alpha_2 = 47.5\%$, $\alpha_3 = 2.5\%$ and $\alpha_4 = 47.5\%$. The path loss exponent is set to be $n = 3$.

Fig. 2 shows the achievable ergodic sum rate, in bps/Hz, versus the transmit SNR, $\rho_t = P_t/\sigma^2$ following the two-phase NOMA scheme explained earlier. In order to achieve the channel conditions $\lambda_{SR1} > \lambda_{SR2}$ and $\lambda_{R2D} > \lambda_{R1D}$, it is assumed that $\lambda_{SR1} = 10$, $\lambda_{SR2} = 3$, $\lambda_{R2D} = 9$ and $\lambda_{R1D} = 5$. As expected, increasing the transmit SNR results in an increase in the overall achievable ergodic data rate. As can be seen in Fig. 2, analytical results are in agreement with the Monte Carlo simulation results at high SNR values, i.e., for $\rho_t > 20$ dB. The slight difference between the analytical results and the simulation results at low SNR values is a consequence of using the asymptotic expression in (36).

In Fig. 3, the effect of the positions of the selected relays, R_1 and R_2 are examined. Three placement scenarios are assessed, with the sum of the distance of the source-to-relay and the distance of the relay-to-destination link fixed at $d_t = d_{SR_i} + d_{R_iD} = 4$ m. In order to maintain the diamond relay topology, it is also assumed that d_{SR1} is 1 m less than d_{SR2} . In the first scenario, both relays are closer to the source than to the destination with $d_{SR1} = 1$ m and $d_{SR2} = 2$ m. In the second scenario, both relays are close to the midpoint between the source and the destination with $d_{SR1} = 1.8$ m and $d_{SR2} = 2.8$ m. Finally, the last scenario, in which $d_{SR1} = 2.4$ m and

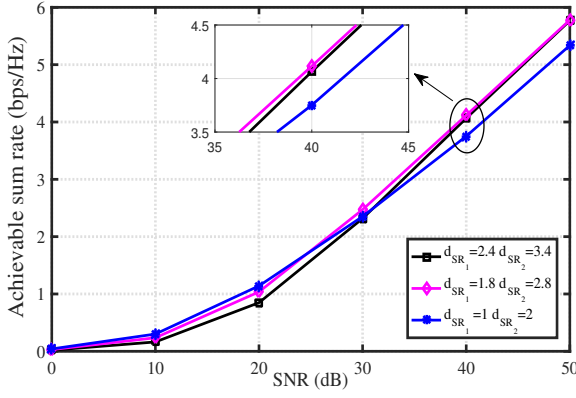


Fig. 3. Achievable sum rate for three test cases with different relay positions, d_{SR_1} and d_{SR_2} .

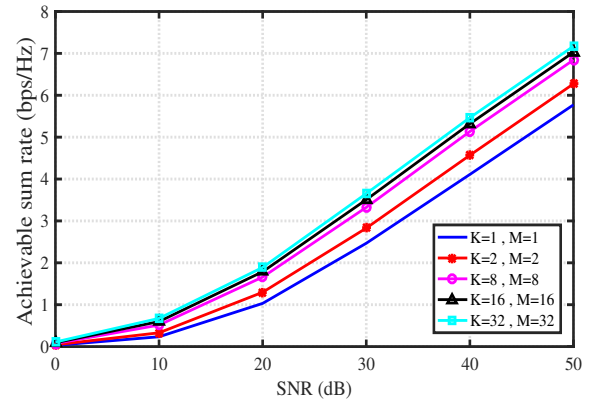


Fig. 5. The achievable sum rate versus the transmit SNR, ρ_t , for various number of relays, $K = M = 1, 2, 4, 8, 16$ and 32 .

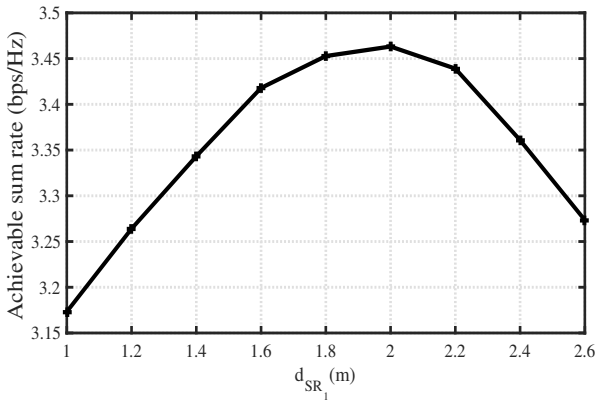


Fig. 4. Achievable sum rate versus various locations of the relay R_1 .

$d_{SR_2} = 3.4$ m, represents a case when both relays are nearer to the destination than to the source.

As can be seen from Fig. 3, the scenario in which both relays are near the midpoint between the source and the destination achieves a higher sum data rate than the other two scenarios. This occurs because the end-to-end capacity is dominated by the minimum of the source-to-relay and the relay-to-destination link capacities, as shown in (15). Thus, while a shorter communication link increases the link's capacity, a longer link has a dominantly negative effect on end-to-end capacity. It should be concluded that both R_1 and R_2 better be as close as possible to the midpoint between the source and the destination. It can be then conjectured that there are optimum locations for relays R_1 and R_2 that result in the maximum possible sum ergodic capacity. Fig. 4 shows the achievable sum rate for different values of the distance, d_{SR_1} . The figure is generated for $\rho_t = 36$ dB and a total distance $d_t = d_{SR_1} + d_{R_1D}$ of 4 m. Results show that there is an optimum location of R_1 for which a maximum sum rate can be achieved. It can be seen from Fig. 4 that for the simulated example, the optimum capacity is achieved by $d_{SR_1} = 1.8 \sim 2$ m.

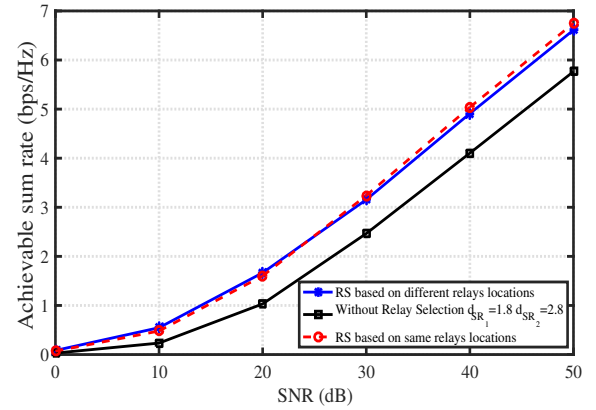


Fig. 6. Achievable data rate with relay selection.

In the next simulation example, the benefit of relay selection and the effect of the number of relays in the selection pools are investigated. In Fig. 5, a comparison of the achievable sum rate based on the number of available relays is introduced. These relays are clustered relatively close together towards the midpoint between the source and the destination with $d_{SR_1} = 1.8$ m and $d_{SR_2} = 2.8$ m. The pair of relays that achieve the highest ergodic capacity is selected according to (16). As can be observed, the higher the number of relays, the higher the achievable capacity. It can also be observed that while the achievable capacity improves with the number of relays in the selection pool, the amount of improvement diminishes for a larger number of relays. For example, at $\rho_t = 30$ dB, the total sum rate achieved is 2.467 bps/Hz for the case when no relay selection is employed. When increasing the number of relays to $K = M = 16$ relays, the overall sum rate increases to 3.507 bps/Hz, representing about 41% improvement in the achievable sum rate. On the other hand, when the number of relays increases to 32, the achievable sum rate is 3.653 bps/Hz, a modest improvement compared to the case of 16 relays.

In Fig. 6, a more practical scenario is investigated. In this scenario, the available relays are randomly dispersed at various

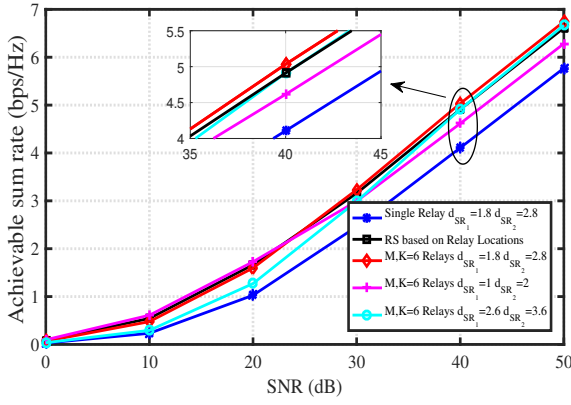


Fig. 7. Comparison between increasing no. of relays vs selecting from different relay locations

locations, and the relay selection scheme selects the optimal relay location. It can be seen from Fig. 6 that at SNR $\rho_t = 30$ dB, the achievable sum rate for a single pair of relays located at the optimum location between the source and the destination is 2.4 bps/Hz. However, when relay selection is performed from six pairs of relays randomly deployed at various locations in the range $d_{SR1} \in [1.1 : 2]$, the achievable sum rate increases to 3.1 bps/Hz. Similarly, if six pairs of relays are all located around the optimum location, the improvement in the data rate is the same as that achieved by randomly deployed relays. It can be concluded then that to benefit from relay selection, it is not necessary to have all relays located at the optimum location, and it is sufficient to use randomly deployed relays.

The previous observations are emphasized by the results in Fig. 7, in which multiple cases of relay selection are compared to each other. It can be observed that relay selection from a pool of relays located at the optimum location results in the highest achievable sum rate. In addition, randomly deployed relays results in better performance than relay selection from relays located all at a location other than the optimum location. In practical systems, randomly deployed relayed can be used with relay selection to maximize the benefit of NOMA-based diamond relay networks.

V. CONCLUSION

This paper studied a NOMA-based diamond relaying network model with a single source, single destination, and multiple randomly deployed decode-and-forward relaying nodes. Motivated by the benefits of non-orthogonal multiple access and relay selection, the considered model can result in higher achievable data rates than those achieved by each alone. The sum data rate achieved by the proposed system was derived, and design guidelines for optimum relay selection policy were recommended. Simulation results showed that higher data rates are achieved when the two selected relays are such that one is near to the source, and the other is near to the destination. It was also shown that the larger the number of relays from which a pair is selected, the higher the achievable data rate.

The performance gain obtained from increasing the number of relays at an optimum location has also been analyzed. It was shown that such gains diminish with a higher number of relays. Further extension of this work will include the joint optimization of the relay placement and power allocation of the selected relays in randomly deployed NOMA-based NDR networks.

REFERENCES

- [1] M. Z. Chowdhury, S. Ahmed, and Y. M. I. N. Jang, "6G Wireless communication systems : Applications, requirements, technologies, challenges, and research directions," in *IEEE Open Journal of the Communications Society*, vol. 1, pp. 957–975, 2020.
- [2] L. Chettri and R. Bera, "A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 16–32, 2020.
- [3] F. Guo, F. R. Yu, H. Zhang, X. Li, H. Ji, and V. C. M. Leung, "Enabling massive iot toward 6g: A comprehensive survey," *IEEE Internet of Things Journal*, vol. 8, no. 15, pp. 11 891–11 915, 2021.
- [4] M. Xu, F. Ji, M. Wen, and W. Duan, "Novel receiver design for the cooperative relaying system with non-orthogonal multiple access," *IEEE Communications Letters*, vol. 20, no. 8, pp. 1679–1682, 2016.
- [5] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A survey of non-orthogonal multiple access for 5G," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 3, pp. 2294–2323, 2018.
- [6] I. Budhiraja, N. Kumar, S. Tyagi, S. Tanwar, Z. Han, D. Y. Suh, and M. J. Piran, "A systematic review on NOMA variants for 5G and beyond," *IEEE Access*, vol. 9, pp. 85573–85644, 2021.
- [7] M. Al-Imari, P. Xiao, M. A. Imran, and R. Tafazolli, "Uplink non-orthogonal multiple access for 5G wireless networks," *11th International Symposium on Wireless Communications Systems (ISWCS)*, pp. 781–785, 2014.
- [8] M. Zeng, W. Hao, O. A. Dobre, and Z. Ding, "Cooperative NOMA: state of the art, key techniques, and open challenges," *IEEE Network*, vol. 34, no. 5, pp. 205–211, 2020.
- [9] M. Liaqat, K. A. Noordin, T. Abdul Latif, and K. Dimyati, "Power-domain non-orthogonal multiple access (PD-NOMA) in cooperative networks: An overview," *Wireless Networks*, vol. 26, no. 1, pp. 181–203, 2020.
- [10] Z. Ding, H. Dai, and H. V. Poor, "Relay selection for cooperative NOMA," *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 416–419, 2016.
- [11] Z. Yang, Z. Ding, Y. Wu, and P. Fan, "Novel relay selection strategies for cooperative NOMA," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 11, pp. 10 114–10 123, 2017.
- [12] Thi Nguyen, TT., Do, DT., Chen, YC. et al. "New look on relay selection strategies for full-duplex multiple-relay NOMA over Nakagami-m fading channels," *Wireless Networks*, vol. 27, pp. 3827–3843, 2021.
- [13] D. T. Do, M. S. Van Nguyen, T. A. Hoang, and M. Voznak, "NOMA-assisted multiple access scheme for IoT deployment: Relay selection model and secrecy performance improvement," *Sensors journal - Internet of Things and Machine-to-Machine Communication Issue*, vol. 19, no. 3, 2019.
- [14] D. Wan, M. Wen, H. Yu, Y. Liu, F. Ji, and F. Chen, "Non-orthogonal multiple access for dual-hop decode-and-forward relaying," *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2016.
- [15] D. Wan, M. Wen, F. Ji, H. Yu, and F. Chen, "On the achievable sum-rate of NOMA-based diamond relay networks," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1472–1486, 2019.
- [16] B.-Y. Huang, Y. Lee, and S.-I. Sou, "Joint power allocation for NOMA-based diamond relay networks with and without cooperation," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 428–443, 2020.
- [17] Y. Lee, B. Y. Huang, and S. I. Sou, "An efficient algorithm for joint power allocation in NOMA-based diamond relay networks," *International Computer Symposium (ICS)*, pp. 288–293, 2020.
- [18] Y. Zhang, Z. Yang, Y. Feng, and S. Yan, "Performance analysis of cooperative relaying systems with power-domain non-orthogonal multiple access," *IEEE Access*, vol. 6, pp. 39 839–39 848, 2018.
- [19] D. Zwillinger, *Table of Integrals, Series, and Products: Eighth Edition*, 2014.