A Lightweight Machine Learning Assisted Power Optimization for Minimum Error in NOMA-CRS over Nakagami-m channels

Ferdi Kara, Senior Member, IEEE, Hakan Kaya, Halim Yanikomeroglu, Fellow, IEEE.

Abstract—Non-orthogonal multiple access based cooperative relaying system (NOMA-CRS) has been proposed to alleviate the decay in spectral efficiency of the conventional CRS. However, existing NOMA-CRS studies assume perfect successive interference canceler at the relay and mostly investigate sum rate whereas the error performance has not been taken into consideration. In this paper, we analyze error performance of the NOMA-CRS and the closed-form bit error probability (BEP) expression is derived over Nakagami-m fading channels. Then, thanks to the high performance of machine learning (ML) in challenging optimization problems, a joint power sharing-power allocation (PS-PA) scheme is proposed to minimize the bit error rate (BER) of the NOMA-CRS. The proposed ML-assisted optimization has a very low online implementation complexity. Based on provided extensive simulations, theoretical BEP analysis is validated. Besides, the proposed ML-aided PS-PA provides minimum BER (MBER) and outperforms previous PA strategies for the NOMA-CRS notably.

Index Terms—error analysis, cooperative relaying, NOMA, optimum power allocation, machine learning, Nakagami-m fading

I. INTRODUCTION

Non-orthogonal multiple access (NOMA) has been envisioned as a key technology for the future wireless networks due to its high spectral efficiency [1]; therefore, its integration into other physical techniques has been widely investigated such as free-space optics [2] and space-time-block coded schemes [3]. NOMA-based cooperative relaying system (CRS) is one of the most attracted topics since the spectral inefficiency in conventional CRS can be eliminated thanks to NOMA integration [4]. The ergodic capacity of the NOMA-CRS is analyzed and it is shown that NOMA-CRS has a better capacity performance than the conventional CRS over different fading channels [4], [5]. This performance enhancement in capacity has led researches to investigate NOMA-CRS and various NOMA-CRS schemes have been proposed/analyzed in terms of capacity and outage performances [6]–[9]. Xu et al [6] propose a novel receiver design and prove that the capacity gain of the NOMA-CRS can be further improved. Then, Zhang et al [7] have analyzed the capacity and outage performances of the NOMA-CRS for different transmission strategies under imperfect channel state information (CSI). Then, it is assumed to be two relays in the network and ergodic capacity is investigated [8]. Moreover, Abbasi et al [9] consider an amplify-forward (AF) relay in NOMA-CRS and provide an approximate expression for ergodic rate. However, existing studies mostly assume that perfect successive interference canceler (SIC) is implemented at the relay which is quite strict/unreasonable assumption and should be relaxed due to the nature of wireless communications. Besides, all in previous works, the analysis is based on only the SINR definitions which do not represent the performance when an actual modulator and/or demodulator (e.g., SIC) are implemented. In addition, in all previous studies [4]–[9], NOMA-CRS has been analyzed only in terms of informational-theoretic perspectives (i.e., capacity and outage) whereas only in [10], [11], the approximate bit error probability (BEP) has been conducted for only two subsets of the NOMA-CRS. Besides, these papers [10], [11] consider only Rayleigh fading channels. However, to the best of the authors knowledge, the exact BEP of the NOMA-CRS has not been derived, yet, although, it is one of the most important key performance indicators (KPIs).

On the other hand, machine learning (ML) techniques have been proved to be efficient alternatives in solving challenging wireless communications problems and have started to attract great recent attention from the communication society [12]. Thus, in this paper, we take the advantage of ML in power optimization to minimize the error performance of the NOMA-CRS. The main contributions are as follow:

• We derive the exact BEP of the NOMA-CRS in closed-form and to the best of the authors’ knowledge, this is the first study which investigates the error performance of the NOMA-CRS with the imperfect SIC, a realistic scenario. In addition, this paper considers Nakagami-m fading channels which represent more comprehensive channel conditions than Rayleigh fading channels. Theoretical analysis is validated via computer simulations.

• We propose a lightweight machine learning (ML)-aided joint power sharing-power allocation (PS-PA) optimization for the NOMA-CRS under the minimum bit error rate (MBER) constraint. This novel solution proposes an optimization not only the first in terms of error performance but also the first joint optimization in terms of any performance metric. Based on extensive simulations, proposed ML-aided PS-PA has performed well in predicting actual optimal values, obtained by an exhaustive search, and it outperforms previous PA strategies in terms of error performance with a very low complexity.

The rest of the paper is organized as follows. In Section II, NOMA-CRS is defined. Then, Section III provides the theoretical BEP analysis. ML-aided optimum PS-PA is introduced in Section IV. In Section V, simulation results are presented. Finally, Section VI concludes the paper with discussions.

II. SYSTEM MODEL

A NOMA-CRS where a source (S) is willing to reach out the destination (D) and a half-duplex relay (R) helps for it [4], [5], [7] is considered. All nodes are assumed to be equipped with single antenna and the flat fading channel coefficient between each nodes (i.e., $h_{\lambda}$, $\lambda = sr, sd, rd$) follows Nakagami-m distribution with shape $m_{\lambda}$ and spread $\Omega_{\lambda}$ parameters. In order to overcome the inefficiency of the conventional CRS of device-to-device communication, NOMA is applied for two intended/consecutive symbols of the destination in the first
phase of the communication. Then, this total superposition-coded symbol is conveyed to both the destination and the relay, hence the received signals in the first phase are given as
\[
y_{\lambda} = \sqrt{P_s} \left( \sqrt{\alpha}x_1 + \sqrt{1-\alpha}x_2 \right) h_\lambda + w_\lambda, \quad \lambda = sr, sd,
\]
where \(P_s\) is the transmit power of the source. \(\alpha\) is the power allocation (PA) coefficient. \(x_1\) and \(x_2\) are the two symbols of the destination to be transmitted simultaneously in NOMA-CRS -they are transmitted sequentially in conventional CRS-. \(w_\lambda\) denotes the additive white Gaussian noise (AWGN) with \(N_0\) variance. \(\alpha < 0.5\) is assumed. Thus, in the first phase, both relay and destination detect \(x_2\) symbols by treating \(x_1\) symbols as noise. Then, the relay implements SIC to detect \(x_1\) symbols and forwards detected \(\hat{x}_1\) symbols to the destination in the second phase [4, 5, 7]. The received signal by the destination in the second phase is given as
\[
y_{rd} = \sqrt{P_r} x_1 h_{sr,d} + w_{rd},
\]
where \(P_r\) is the transmit power of the relay. Finally, the destination detects \(x_1\) symbols based on \(y_{rd}\).

III. Bit Error Probability (BEP) Analysis

In order to derive total BEP of the NOMA-CRS, BEPs for two symbols should firstly be derived and averaged. Thus, the average BEP (ABEP) of the NOMA-CRS is given by
\[
P(e_{2e}) = \frac{P_{x_1}(e) + P_{x_2}(e)}{2},
\]
where \(P_{x_1}(e)\) and \(P_{x_2}(e)\) denote the BEPs of \(x_1\) and \(x_2\) symbols, respectively.

Since the \(x_2\) symbols are conveyed to the destination only in the first phase and the \(x_1\) symbols are treated as noise in detection, the BEP for \(x_2\) symbols will be the same with BEP of far user in downlink NOMA. The conditional BEP of far user’s symbols in NOMA schemes is given as
\[
P_{x_2}(e|\gamma_{sd}) = \sum_{i=1}^{N} \zeta_i Q \left( 2\nu_i \rho_{s} \gamma_{sd} \right),
\]
where \(\gamma_{sd} = |h_{sd}|^2\) and \(\rho_s = P_s/N_0\) are defined. \(N\), \(\zeta_i\) and \(\nu_i\) coefficients change according to chosen modulation constellation pairs for \(x_1\) and \(x_2\) symbols. In case BPSK is used for both symbols (i.e., \(x_1\) and \(x_2\)), by following steps [11] Eq.(3)-(4), it is derived that \(N = 2\), \(\zeta_i = 0.5\) and \(\nu_i = 1 \mp 2\sqrt{\alpha - \alpha^2}\).

Proof: In NOMA schemes, the BEP highly depends on chosen constellation pairs and we should consider the signal energy in each scenario (i.e., that is not fixed due to superimposed symbols). Considering the baseband symbol \(x_i = \mp 1\) for BPSK, the received signal at the nodes are given in Table I along with the correct decision rule. For instance, when \(b_1 b_2 = 00\) bit-stream is conveyed for \(x_1\) and \(x_2\) symbols, the received signal at the nodes becomes \(y_{\lambda} = ( -\sqrt{\alpha} - \sqrt{1-\alpha}) h_\lambda + n_\lambda\). To detect \(x_2\) symbols correctly, the received signal, \(y_{\lambda} < 0\) should be satisfied; thus, the erroneous detection probability is given by
\[
P \left( n_\lambda \geq \left( \sqrt{\alpha} + \sqrt{1-\alpha} \right) h_\lambda \right).
\]
By repeating all scenarios with the given decision rules, we obtain coefficients as given above, so the proof is completed.

Then, by averaging the conditional BEP over instantaneous \(\gamma_{sd}\) (it follows Gamma distribution), with the aid of [13], the BEP of \(x_2\) symbols is derived as
\[
P_{x_2}(e) = \left\{ \begin{array}{ll}
\sum_{k=1}^{2} \frac{1}{4} \left[ 1 - \mu^2(b_k) \right] \left( \frac{2}{1} \right) \left( \frac{1-\mu(b_k)}{4} \right), & m_{sd} : integer, \\
\sum_{k=1}^{2} \frac{1}{4} \sqrt{\frac{b_k}{1+b_k}}^{m_{sd}+m_{rd}+0.5} \frac{1}{\Gamma(m_{sd}+1)} \\
\sum_{k=1}^{2} \frac{1}{4} \sqrt{\frac{b_k}{1+b_k}}^{m_{sd}+0.5} \frac{1}{\Gamma(m_{sd}+1)} \\
\end{array} \right.
\]
where \(b_k \triangleq \frac{\rho_s \sqrt{\alpha}}{m_{sd}}\) and \(\mu(z) = \frac{\sqrt{2\alpha} - \sqrt{1-\alpha}}{4}\) are defined. In (5), \(\Gamma(.)\) and \(\mathbb{F}_1(\cdot;\cdot;\cdot)\) denote Gamma [14] Eq. (8.31]) and Gauss Hyper-geometric [14] Eq. (9.10)] functions, respectively.

On the other hand, the \(x_1\) symbols are detected at the relay in the first phase and forwarded to the destination in the second phase. Since the erroneous detection in two phases are statistically independent, with the law of total probability, the BEP of \(x_1\) symbols is given as
\[
P_{x_1}(e) = P_{x_1}^{(sr)}(e) \left( 1 - P_{x_1}^{(rd)}(e) \right) + \left( 1 - P_{x_1}^{(sr)}(e) \right) P_{x_1}^{(rd)}(e),
\]
where \(P_{x_1}^{(sr)}(e)\) and \(P_{x_1}^{(rd)}(e)\) denote the BEPs of \(x_1\) symbols between nodes S-R (i.e., first phase) and R-D (i.e., second phase), respectively.

The BEP of \(x_1\) in the second phase (i.e., \(P_{x_1}^{(rd)}(e)\)) can easily be obtained, since no interference is encountered in the second phase (only transmission from R-D exists). Thus, the conditional BEP of \(x_1\) in the second phase turns out to be the well-known error probability over fading channels. For BPSK, it is given as
\[
P_{x_1}^{(rd)}(e|\gamma_{rd}) = Q \left( \sqrt{2\rho_r} \gamma_{rd} \right) \text{ where } \rho_r = \frac{P_r}{N_0}.
\]

| Bit-stream \((b_1 b_2)\) | Received (Remained) Signal | Correct Decision rule |
|----------------------|-----------------------------|-----------------------|
| 00                   | \(-\sqrt{\alpha} - \sqrt{1-\alpha}) h_\lambda + n_\lambda\) | \(y_{\lambda} < 0\) |
| 01                   | \(-\sqrt{\alpha} + \sqrt{1-\alpha}) h_\lambda + n_\lambda\) | \(y_{\lambda} \geq 0\) |
| 10                   | \(\sqrt{\alpha} + \sqrt{1-\alpha}) h_\lambda + n_\lambda\) | \(y_{\lambda} < 0\) |
| 11                   | \(\sqrt{\alpha} - \sqrt{1-\alpha}) h_\lambda + n_\lambda\) | \(y_{\lambda} \geq 0\) |

\(\alpha\) For the representation simplicity, we do not represent channel coefficient and AWGN.
The BEP over Nakagami-m fading channels is given as \[ P_{x_1}(e) = \left\{ \begin{array}{ll} \frac{1}{2} - 1 - \mu^2(p) \sum_{i=0}^{m_r-1} \left( \frac{1}{i!} \mu^2(p) \right)^i, & m_r : \text{integer}, \\
\frac{1}{\sqrt{\pi} (1+p)^{\frac{m_r-1}{2}} \Gamma(m_r+0.5)^{\frac{1}{2}}} \times 2F_1 \left( 1, m_r + 0.5; m_r + 1; \frac{1}{1+p} \right), & m_r : \text{non-integer}, \end{array} \right. \] \] (7)

where \( p \triangleq \frac{\rho_m}{m_r} \) is defined.

However, in order to derive the BEP in the first phase (i.e., \( P_{x_1}(e) \)), much more effort is required. In the first phase, since the superimposed signal is received by the relay, the SIC should be implemented and the error propagation during SIC should be taken into consideration in the BEP analysis. The conditional BEP of \( x_1 \) symbols (i.e., near user in conventional downlink NOMA) can be given in the form

\[ P_{x_1}(e|\gamma_{sr}) = \sum_{i=1}^{L} \eta_i Q \left( \sqrt{2\vartheta_i \rho_s \gamma_{sr}} \right), \] \] (8)

where \( L, \eta_i \) and \( \vartheta_i \) change according to the modulation pairs. For BPSK, when we repeat the steps [11] Eq.(6)-(9), we obtain \( L = 5, \eta = \{2,1,1,1,1\} \) and \( \vartheta = \{1,1,1,1,1\} \), and the decision rule in this case, the conditional probability rule, it becomes

\[ 2F_1 \left( 1, m_r + 0.5; m_r + 1; \frac{1}{1+e} \right), m_r : \text{non-integer}, \]

where \( e \triangleq \frac{\rho_m}{m_r} \) is defined. Finally, by substituting (7) and (9) into (6) and then by substituting (5) and (6) into (3), the ABEP of the NOMA-CRS is derived in the closed-form as (10) (see the bottom of the page).

### IV. POWER OPTIMIZATION FOR MBER

Considering the total power consumption at the nodes, let us define a PS coefficient (i.e., \( \beta \)). The total power \( (P_T) \) is shared as \( P_s = \beta P_r \) and \( P_c = (1 - \beta) P_T \). The joint PS-PA optimization is defined as the \( \alpha, \beta \) pair which minimizes the ABEP of the NOMA-CRS. Thus, it is given as

\[ [\alpha^*, \beta^*] = \arg \min_{\alpha, \beta} P_{e}(e). \] \] (11)

And, it is derived by solving

\[ [\alpha^*, \beta^*] = \left( \frac{\partial^2 P_{e}(e)}{\partial \alpha \partial \beta} = 0 \right). \] \] (12)

To the best of the authors’ knowledge, (12) cannot be analytically solved in closed-form. It can be obtained by iterative algorithms such as a brute-force/full-search algorithm [15] by computing the ABEP values for all PS-PA coefficients to obtain the minimum. However, this costs a high computational complexity in the online implementation thereby increasing latency and is inappropriate for practical implementations. Besides, with this high computational operation, high power consumption will be also required at the relay which yields an unfairness for the relay. Thus, we propose an ML-based model to obtain the optimum PS-PA pair for any condition in NOMA-CRS. The proposed ML network is trained offline to reduce computation time/complexity/latency and it is implemented online as being in all ML-based solutions in physical layer communications [12], [16]. Therefore, the computational complexity for solving the optimization problem is shifted to the offline training stage and a very low online implementation complexity is achieved.

### A. ML-aided Optimum PS-PA

1) Proposed ML Model: We built a three-layered (i.e., input, hidden, and output layers) fully connected neural network (NN) model to compute the optimum PS-PA pair for the MBER in NOMA-CRS. Hence, we redefined the PS-PA optimization problem

\[ [\alpha^*, \beta^*] = f(m_\lambda, \Omega_\lambda). \] \] (13)

According to (13), the built NN model has 6 inputs (neurons) and 2 outputs (neurons). The number of neurons in hidden
layer is 10 and \textit{Levenberg-Marquardt} is used for the learning algorithm. Minimum performance gradient, learning (mu) decrease factor and learning (mu) increase factor are set as $1e-7$, 0.1 and 10, respectively. The network parameters are empirically determined such that increasing number of layers and/or neurons do not provide a remarkable performance gain and the training performance converges. Therefore, not to increase the online implementation complexity, we came up with that three-layered NN is enough to solve the problem defined in (11) although it can be built with much deeper networks (more hidden layers).

2) Training and Testing: To train NN model, we firstly create a dataset for different channel and power conditions. Then, we train the NN model to predict the optimum PS-PA pair for given training inputs with minimum MSE compared to desired training outputs. The dataset generation and training algorithm is given in Algorithm 1. In Algorithm 1, we divide

the dataset into two groups: 90% for training and 10% for test datasets. We train the network with training dataset to minimize the MSE where we use a cease criterion that we stop training if the MSE is not improved at least by $e-5$ within two consecutive epochs. Then, we test the trained network with the test dataset (i.e., which is not used in training stage) and re-train the network until we obtain a good (convincing) test performance. At the end of this re-training procedure, we have concluded with the best performance metrics as obtained $3.43e-4$ MSE and 0.9984 and regression for training and $3.44e-4$ MSE and 0.9842 regression for testing compared to desired (numerically calculated) optimum PS-PA pairs.

\textbf{Algorithm 1 Dataset Generation and Training}

1: Data Set Generation
2: \textbf{Multiple For Loops} ($\Omega_L = [0:1:10]$, $\lambda_L = [0.5:0.5:4]$ and $\rho_T = P_T/N_0 = [0:5:20]$ dB)
3: Solve (12) by numerical tools, then, label the obtained results as desired [$\alpha^*, \beta^*$] outputs for the given inputs
4: \textbf{End For Loops}
5: Portion the dataset 90% for training and 10% for testing
6: \textbf{While}
7: Train the network with the training inputs (90%) to minimize the MSE
8: Test the trained network for the testing inputs (10%)
9: \textbf{Unless} the test performance satisfies

\begin{align*}
P(c|e) &= \begin{cases} 
\frac{2}{3} \sum_{k=0}^{m-1} \left[ -\frac{\mu^2(b_k)}{4} \right] + \sum_{k=1}^{m-1} \left[ -\frac{\mu^2(c_k)}{4} \right] \\
\times \left[ 1 - \frac{1}{2} \left( 1 - \frac{\mu^2(p)}{4} \right) \right] \\
\times \left[ 1 - \frac{1}{2} \left( 1 - \frac{\mu^2(p)}{4} \right) \right] \\
\times \left[ 1 - \frac{1}{2} \left( 1 - \frac{\mu^2(p)}{4} \right) \right] \\
\times \left[ 1 - \frac{1}{2} \left( 1 - \frac{\mu^2(p)}{4} \right) \right]
\end{cases} 
\quad \forall \lambda, \ m_{\lambda} : \text{integer},
\end{align*}

In the ML-based solutions, the computational complexity is related to the online implementation. Therefore, we focus on the online implementation complexity (feed-forward calculation). The online implementation complexity is just $O(104)$ which consists of 60 and 20 weight multiplications from-input-to-hidden and from-hidden-to-output layers, respectively. Total 12 adds for biases and 12 activation functions (i.e., \textit{tansig} function) are computed on hidden and output layers neurons.

On the other hand, to optimize PS-PA by an iterative algorithm (i.e., brute-force/full-search algorithm), the ABEP expression (10) should be computed for all PA-PS pairs and compared with each other. Thus, by considering arithmetic/logical operations in (10) (includes computing high-complex $\Gamma(.)$ and $2F_1(\cdot;\cdot)$ functions), the complexity of the full search algorithm is obtained as $O(M^2(2m_{rd} + 5m_{sr} + m_{rd} + 96) + 1)$ for non-integer case where $1/M$ denotes the step size (resolution in search) for PS and PA. For even all integer case, it reduces to $O(M^2(2m_{rd} + 5m_{sr} + m_{rd} + 9) + 1)$. Besides, this computational complexity should be repeated whenever the channel conditions change. However, the proposed model has only $O(104)$ online complexity. Therefore, the complexity (a.k.a. latency) is reduced, and the optimization is implementable in practical scenarios. Since we propose a joint PS-PA optimization, this should be computed at both source and relay. Therefore, reducing the complexity is essential. With the proposed method, we shifted the complexity to the offline training stage (at a server with high computational capacity); thus, the online implementation complexity at the source and relay is limited (e.g., centralized offline training and distributed online implementation). This is also very important for extended scenarios such as multi-relay schemes where the same online complexity will be required even if a relay selection is applied \cite{17}.
V. Numerical Results

The Fig. 1 presents the BER of the NOMA-CRS for different shape and spread parameters with respect to \( \rho_r \). In Fig. 1.a, it is assumed that \( m_\lambda \) is equal \( \forall \lambda \). The outcomes of the online implementation of the proposed ML network are used as PS-PA strategy. It is noteworthy that the theoretical analysis matches perfectly with the simulations. As expected, the shape parameter \( (m_\lambda) \) denotes the diversity order (superscript of the \( 1/\text{SNR} \) when \( \text{SNR} \to \infty \)) while increasing the spread parameter \( (m_\lambda) \) provides only a horizontal gain in error performance. Then, to reveal the effect of the shape parameter for each node, we present the BER of the NOMA-CRS in Fig. 1.b when \( m_\lambda \) is not equal \( \forall \lambda \). Likewise in conventional CRS, the diversity order of the NOMA-CRS is driven by the lowest shape parameter (i.e., \( \min \{m_{sr}, m_{sd}, m_{rd}\} \)). Nevertheless, due to the cooperative communication, an error propagation occurs from the relay to the destination, the full diversity cannot be observed. Hence, to achieve the full diversity, by considering the total consumed power, the PS and PA should be jointly optimized for MBER. For instance, when \( m_{sr} = m_{sd} = 4 \) and \( m_{rd} = 2 \), the diversity order of the NOMA-CRS is equal to 2. Nevertheless, the other \( m \) parameters have also an effect on the error performance, especially in the high SNR region. With lower \( m \) parameter between S-R (e.g., \( m_{sd} = m_{rd} = 4, m_{sr} = 2 \)), the \( x_1 \) symbols in the first phase has a poor error performance, thus it causes an error propagation from relay-to-source and the error performance gets worse. It can be easy seen in high SNR regime although this case has also the diversity order of 2. Since the total BER is the average of two symbols, the BER of NOMA-CRS highly depends on their individual BEPs. Hence, we should guarantee that none of them pulls down the BER. Then, to reveal the effectiveness of the proposed PS-PA scheme, we present the BER of the NOMA-CRS for the proposed ML-aided PS-PA, fixed PS-PA, and full-search PS-PA in Fig. 1.c. The fixed PS-PA strategy is \( \beta = 0.5 \) (i.e., \( P_s = P_r \)) and \( \alpha = 0.2 \) as assumed in previous NOMA-CRS studies [4–9]. One can easily see that the proposed ML-aided PS-PA performs the same with the full search PS-PA and it outperforms the fixed PS-PA significantly. Hereby, we should note that the online complexity of the proposed ML-aided PS-PA is much less than the full search PS-PA as proved in Section IV.B. In addition, this performance gain over the fixed PS-PA becomes greater with the increase of shape \( (m_\lambda) \) and/or spread \( (\Omega_\lambda) \) parameters.

We present PS-PA comparisons between the outcomes of the proposed NN network and the results obtained by the full search algorithm for various scenarios in Fig. 2. The predictions of the proposed NN network for PS-PA scheme are very close to the full search algorithm. Furthermore, the PS-PA scheme is highly dependant on the relay position. To represent this, we hereby assume that the linear sum of the spread parameters through the link S-R-D is constant (i.e., \( \Omega_{rd} + \Omega_{sr} = 10 \)), thus \( \Omega_{rd} = 10 - \Omega_{sr} \) which can call the relay position (when \( \Omega_{sr} > \Omega_{rd} \), relay is close to source vice versa). It is clearly seen that the optimum PS-PA values get larger when the channel quality between S-R (i.e., \( \Omega_{sr} \)) decreases. It can be explained as follows. If the relay detects \( x_1 \) symbol erroneously, an error propagation occurs from the relay to the destination and this pulls down the total error performance. Hence, not to cause an error propagation, most of the power should be transferred to \( x_1 \) symbols (higher \( \alpha \) and \( \beta \)). However, increasing \( \alpha \) and/or \( \beta \) too much still causes an error propagation due to the SIC operation at the relay. Thus, they should be limited.

Fig. 3 presents BER comparisons between the proposed PS-PA and the optimum PA (in terms of capacity) in [4] with the
fixed PS. In Fig. 3.a, since a joint PS-PA is not implemented in [4], an error propagation occurs in the high SNR region, whereas with the proposed PS-PA, this problem has been resolved. Besides, to emphasize the effect of PS-PA, Fig. 3.b shows the error performance of the NOMA-CRS with respect to $\alpha$ and $\beta$. NOMA-CRS has the best performance when $0.1 \leq \alpha \leq 0.2$ and $0.7 \leq \beta \leq 0.9$. Increasing/decreasing one/b both of them too much causes severe performance for $x_1$ and/or $x_2$ symbols so for the NOMA-CRS. In Fig 3.b, we also mark the points when the proposed PS-PA and the PA in [4] are used. In both figures, it is clear that the proposed ML-aided optimum PS-PA provides the minimum BER for NOMA-CRS and is superior to the other PA strategies.

Lastly, to represent the effectiveness of the proposed PS-PA for higher modulation order schemes, we present BER comparisons for the fixed PS-PA (i.e., $\beta = 0.5$, $\alpha = 0.2$) [4–9], full-search PS-PA and the proposed ML-aided PS-PA in Fig. 4 when QPSK is used for both symbols. Since, the theoretical analysis for QPSK has not been derived, yet, the full-search PS-PA is obtained by simulations where we simulate the NOMA-CRS for $100 \times 100$ PS-PA pairs and chose the PS-PA which has minimum BER. As seen in Fig. 4, the proposed ML-aided PS-PA still outperforms the fixed PS-PA and achieves the full-search PS-PA performance. This simulations can be extended for higher M-QAM schemes; however, in M-QAM signaling, we should define an additional PA constraint for a detectable signal design [18].

VI. CONCLUSION

In this paper, we investigate the error performance of the NOMA-CRS considering imperfect SIC and derive the exact ABEP of the NOMA-CRS over Nakagami-m fading channels. Then, to minimize the BER of the NOMA-CRS, we propose an ML-aided optimum PS-PA strategy. Based on extensive simulations, our proposed ML-aided PS-PA strategy is optimal in terms of MBER criteria and provides a remarkable gain compared to existing PA strategies. This study is the first to optimize PS-PA in NOMA schemes under the BER constraint; thus, it can be further extended for other NOMA systems by re-training the proposed network with the dataset of that NOMA scheme which is seen as a future work.

REFERENCES

[1] M. Vaezi, Z. Ding, and H. Vincent Poor, *Multiple Access Techniques for 5G Wireless Networks and Beyond*, M. Vaezi, Z. Ding, and H. V. Poor, Eds. Cham: Springer International Publishing, 2017.
[2] P. Saxena and M. R. Bhatnagar, “1-bit feedback-based beamforming scheme for an uplink FSO-NOMA system with SIC errors,” *Appl. Opt.*, vol. 59, no. 36, pp. 11 279–11 291, Dec 2020.
[3] M. N. Jamal, S. A. Hassan, and D. N. K. Jayakody, “A new approach to cooperative NOMA using distributed space time block coding,” in *IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2017*.
[4] J.-B. Kim and I.-H. Lee, “Capacity Analysis of Cooperative Relaying Systems Using Non-Orthogonal Multiple Access,” *IEEE Commun. Lett.*, vol. 19, no. 11, pp. 1949–1952, Nov 2015.
[5] R. Jiao, L. Dai, J. Zhang, R. Mackenzie, and M. Hao, “On the Performance of NOMA-Based Cooperative Relaying Systems Over Rician Fading Channels,” *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11 409–11 413, Dec. 2017.
[6] M. Xu, F. Ji, M. Wen, and W. Duan, “Novel Receiver Design for the Cooperative Relaying System with Non-Orthogonal Multiple Access,” *IEEE Commun. Lett.*, vol. 20, no. 8, pp. 1679–1682, Aug. 2016.
[7] 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, Jul. 2018.
[8] D. Wan, M. Wen, F. Ji, H. Yu, and F. Chen, “On the achievable sum-rate of NOMA-based diamond relay networks,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 1472–1486, Feb. 2019.
[9] O. Abbasi, A. Ebrahimi, and N. Mokari, “NOMA inspired cooperative relay system using an AF relay,” *IEEE Wirel. Commun. Lett.*, vol. 8, no. 1, pp. 261–264, Feb. 2019.
[10] Q. Li, M. Wen, E. Basar, H. V. Poor, and F. Chen, “Spatial Modulation-Aided Cooperative NOMA: Performance Analysis and Comparative Study,” *IEEE J. Sel. Top. Signal Process.*, vol. 13, no. 3, pp. 715–728, Jun. 2019.
[11] F. Kara and H. Kaya, “Error Probability Analysis of NOMA-based Diamond Relaying Network,” *IEEE Trans. Veh. Technol.*, vol. 69, no. 2, pp. 2280–2285, Feb. 2020.
[12] A. Zappone, M. Di Renzo, and M. Debbah, “Wireless Networks Design in the Era of Deep Learning: Model-Based, AI-Based, or Both?” *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 7331–7376, Oct. 2019.
[13] M. S. Alouini, “A unified approach for calculating error rates of linearly modulated signals over generalized fading channels,” *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1324–1334, Sep. 1999.
[14] I. Gradshteyn and I. Ryzhik, *Table of Integrals, Series, and Products*, 5th ed. San Diego, CA: Academic Press, Feb. 1994.
[15] C. Grosan and A. Abraham, *Intelligent Systems: A modern Approach*. Berlin, Germany: Springer Heidelberg, 2011.
[16] V. X. Thang, S. Chatzinotas, V.-D. Nguyen, T. Hoang, D. N. Nguyen, M. Di Renzo, and B. Ottersten, “Machine Learning-Enabled Joint Antenna Selection and Predecoding Design: From Offline Complexity to Online Performance,” *IEEE Trans. Wireless Commun.*, Early Access, 2021.
[17] F. Kara, H. Kaya, O. Erkaymaz, and F. Ozturk, “Prediction of the optimal threshold value in DF relay selection schemes based on artificial neural networks,” in *IEEE Int. Symp. Innov. Intell. Syst. Appl.*, Aug. 2016.
[18] F. Kara and H. Kaya, “A True Power Allocation Constraint for Non-Orthogonal Multiple Access with M-QAM Signalling,” in *IEEE Microw. Theory Tech. Wirel. Commun.*, Oct. 2020, pp. 7–12.