applied sciences

Article

Beamforming for the Successive Relaying-Based Cooperative Non-Orthogonal Multiple Access Transmission

Duckdong Hwang [1], Sung Sik Nam [2,*], Janghoon Yang [3] and Hyoung-Kyu Song [1]

1 Department of Information and Communication Engineering, Sejong University, Seoul 05006, Republic of Korea; duckdonh@yahoo.com (D.H.); songhk@sejong.ac.kr (H.-K.S.)
2 Department of Electronic Engineering, Gachon University, Seongnam 13120, Republic of Korea
3 Computer Science, Penn State Abington, Philadelphia, PA 19001, USA; jhyang00@hanmail.net or jxy5427@psu.edu
* Correspondence: ssnam@gachon.ac.kr

Abstract: In this paper, a set of cooperative non-orthogonal multiple access (NOMA) transmission schemes is proposed, in which two half-duplex relaying user terminals (UTs) successively deliver the AP signals toward two NOMA UTs. The successive relaying (SR) of two HD relays emulates the full-duplex (FD) relaying transmission such that the whole transmission of one data block is finished in one time slot. Different from FD relaying, where the removal of self-interference (SI) at the relay is challenging, the SR is suffering from cross-relay interference (CRI) between the two HD relays and the handling of this CRI is the key technical challenge. This CRI suppression is implemented by beamforming with multiple antennas at the relaying UTs, while the beamforming also does the role of maximizing the NOMA transmission efficiency. Two beamforming schemes are proposed, where the first method employs semi-definite relaxation (SDR) with four dimensional (4-D) search. Here, the four dimensions come from the four parameters reflecting the CRI levels at the relaying UTs and signal strengths at the NOMA UTs. This scheme alternates until the 4-D search range is reduced and the sum throughput rate saturates. The second method is computationally simpler than the first one since it avoids the SDR and alternation, and it finds the beamformer set in one shot calculation. Both schemes rely on zero forcing of the CRI. These two schemes along with the HDR relay-based cooperative NOMA scheme are compared in the numerical experiments, and the results exhibit their pros and cons in several aspects.

Keywords: cooperative non-orthogonal multiple access; successive relaying; beamformer; semi-definite relaxation

Citation: Hwang, D.; Nam, S.S.; Yang, J.; Song, H.-K. Beamforming for the Successive Relaying-Based Cooperative Non-Orthogonal Multiple Access Transmission. *Appl. Sci.* **2024**, *14*, 3246. https://doi.org/10.3390/app14083246

Academic Editor: John Xiupu Zhang

Received: 24 February 2024
Revised: 1 April 2024
Accepted: 9 April 2024
Published: 12 April 2024

1. Introduction

A huge number of small devices such as the internet of things (IoT) will be served in the coming 6-th generation (6G) wireless networks and consequently more wireless resources and improved latency handling capability of the systems are required. The non-orthogonal multiple access (NOMA) transmission technologies [1] draw research attention in fulfilling these requirements and many research works have been resulted so far. The power domain NOMA [2–4] is attractive since it does not need additional spectral resources or a specific multiple access scheme in accommodating the NOMA principle, where multiple message signals are superposed in a single spectral resource. Since the signals destined to different users are allocated onto separate orthogonal resources in the orthogonal multiple access (OMA), the received signals of the OMA schemes are almost free of the inter-user interference (IUI). This is not true in the power domain NOMA scheme, and the successive interference cancellation (SIC) is a common way to handle the IUI. To maximize the benefit of the NOMA systems, resource allocation acts in a crucial role [5–8], which incurs additional complexity however. The NOMA principle can be applied to millimeter wave band communications as well [9–11].

1.1. Related Works

Multiple antennas and the related technologies can be employed to improve the NOMA system performance in various metrics by exploiting the spatial dimensions [3,4,12–19]. Even the security of the NOMA network can be improved [20] by multiple antenna technologies. In the multiple input multiple output (MIMO) NOMA, both the access point (AP) and the UTs are equipped with multiple antennas, the clustering of UTs can be carried out, the signals pertaining to different clusters can be separated, and the UTs of each cluster form an independent NOMA channel [13–15]. Such useful clustering is not possible in multiple input single output (MISO) NOMA systems, where only the AP is equipped with multiple antennas, and the AP beamforming can serve only single channel NOMA UTs [12,21,22]. Optimal beamforming for this MISO-NOMA system is considered in [21]. Some of the NOMA UTs close to AP can relay the signal for the other UTs far away from the AP in the cooperative NOMA system. The relaying UTs help the AP transmission [23–28] to improve the diversity performance and the power efficiency at the expense of additional spectral resource. The relaying UTs in the cooperative NOMA may operate either in half-duplex (HD) mode or in full-duplex (FD) mode. In HD mode [25–27], the AP signals and the UT relaying signals are transmitted in separate spectral resources, which simplifies the design with extra spectral resources. On the other hand, the FD mode relaying [23,24] is spectrally efficient since the two signal transmissions are overlapped in the same spectral resource, though the handling self-interference (SI) at the relaying UTs becomes more challenging. One method to avoid this challenge and to enjoy the FD advantage of efficient spectral efficiency is adopting successive relaying (SR), where two HD relaying UTs relays the AP signal in alternating fashion [28]. In this way, two SR cooperative relaying UTs can emulate the FD operation and save additional wireless resource without generating SI.

1.2. Contribution

This article considers the beamforming in the SR cooperative NOMA system, where multiple antennas are adopted at the two SR relaying UTs to improve the performance of the cooperative NOMA system. While this SR cooperation avoids SI, it generate cross-relay interference (CRI) between relaying UTs, and hence, the removal of this CRI becomes a new challenge, which is handled by the beamforming proposed in this work. To the author's best knowledge, the beamforming problem for the SR cooperative NOMA system has not been discussed, while it is essential to utilize the spatial dimension to improve transmission efficiency. We present two sub-optimal beamformer design algorithms for the SR cooperative NOMA system since the problem structure does not easily lend itself to the optimal approach. First, the decoding conditions at different NOMA UTs and the CRI levels at the relaying UTs lead us to consider a semi-definite relaxation (SDR)-based beamformer design, where two sets of thresholds on SINRs at UTs and CI power levels are utilized to form sets of constraints for the SDR. The first algorithm finds two sets of beamformer vectors for the two relaying UTs in an alternating fashion, where the considered ranges of thresholds are reduced in every alternation. This alternating search method of multi-dimensional thresholds is hinted from a popular one-dimensional search. In every alternation, the power levels of the found transmit beamformer vectors are compared to the relay transmit power levels, and the ranges of thresholds in the next alternation are determined based on the comparison, and while the relay transmit beamformer vectors are found from the SDR, the receive beamformer vectors are set to zero-force CRI. The first algorithm depends on iterative applications of the SDR algorithm and requires much computation. The second algorithm is a reference scheme with simplified computation without SDR and alternations. The relay UTs receive antennas to try to zero-force the RCI between relaying UTs, while the transmit beamformers are set to the primary UTs in the second hop.

In the numerical experiments, these two beamforming schemes and half-duplex relay (HDR)-based cooperative NOMA systems are compared in their performance in three different environments. The AP power, the relay UT power and the number of relay

UT antennas are used for the experiment parameters to discuss out the pros and cons of proposed schemes.

Section 1 presents the system model of the proposed SR cooperative NOMA. In Section 2, two beamforming algorithms are explained. Section 3 presents numerical results for a few experiments. Finally, Section 4 concludes this paper.

Notations: The notations $\mathbf{A}^H$ and $\mathbf{A}^T$ are the Hermitian transpose and the transpose of a matrix $\mathbf{A}$, respectively. $\mathbf{Tr}[\mathbf{A}]$ takes the trace of the matrix $\mathbf{A}$ and rank$\{\mathbf{A}\}$ returns the rank of the matrix $\mathbf{A}$. $\|\mathbf{a}\|$ denotes the ℓ_2-norm of a vector $\mathbf{a}$. The notation $\mathbf{A}_{i,j}$ represents the element at the i-th row and the j-th column of $\mathbf{A}$. The notation diag$[\mathbf{v}]$ produces a diagonal matrix with the elements of the vector $\mathbf{v}$ on its diagonal. The vector notations $\mathbf{0}_k$ and $\mathbf{1}_j$ represent all zero vector with k elements and all one vector with j elements, respectively. $\mathcal{CN}(\mathbf{0}, \mathbf{C})$ denotes the complex white Gaussian distribution of random vector with zero mean vector $\mathbf{0}$ and the covariance matrix $\mathbf{C}$. $\mathbb{C}^N$ denotes the N dimensional complex vector space. Finally, $\mathbb{E}_n[x]$ takes the expectation of x with respect to n.

2. System Model

Consider the cooperative NOMA system where an AP transmits messages to two NOMA user terminals (U_1 and U_2) through two half-duplex relays (HDRs) (R_1 and R_2) with multiple antennas, where the two relays successively relay the AP signals to emulate full-duplex relaying. In power domain NOMA transmission, the AP transmits a superposed signal toward multiple UTs, where the individual signals for the UTs are added up and sent in the same spectral resource. The UTs apply the SIC decoding principle to decode the particular message destined to themselves. The two HD relays operate with the DF protocol. The AP and the user terminals are equipped with single antennas. In even time slots, R_1 transmits the signal from the AP received in the previous odd time slots, while R_2 helps the AP transmission in odd time slots heard from the previous even time slots. The two relays equipped with M antennas, respectively, use $M \times 1$ receive beamformer vectors $\mathbf{w}_{R,i}, i = 1,2$, with $\|\mathbf{w}_{R,i}\|^2 = 1, i = 1,2$ and use $M \times 1$ transmit beamformer vectors $\mathbf{w}_{T,i}, i = 1,2$, with $\|\mathbf{w}_{T,i}\|^2 = P_r, i = 1,2$ when P_r is the relay transmit power. The decode and forward (DF) protocol is used for the two relays. The multiple-input single-output (MISO) channel from the AP to the relay R_i is denoted by the $M \times 1$ vector $\mathbf{h}_i$, the multiple-input multiple-output (MIMO) channel from R_1 to R_2 is denoted as an $M \times M$ matrix $\mathbf{G}_1$, and the opposite direction channel is denoted by the $M \times M$ matrix $\mathbf{G}_2$. Furthermore, the MISO channel from R_i to U_j is denoted by a $M \times 1$ vector $\mathbf{g}_{i,j}$. Note that all the elements of these channel matrices and vectors are independent and identically distributed random variables with $\mathcal{CN}(0, 1)$. Therefore, we assume here flat fading narrow band wireless channels with Rayleigh distributed channel parameters. Note that more aggressive channel assumptions such as wide band multi-path model or interference from an outside entity may complicate the problem to be solved and we leave such cases for future research topics. We assume that full channel state information (CSI) is available at the AP and relays so that all the beamformer design is conducted at the AP or at the relays.

For the binary variable $i = 1,2$, the notation $\bar{i} = 2$ when $i = 1$ and $\bar{i} = 1$ when $i = 2$. The received signals at the relays and at the users at time slot n are expressed, respectively, as

$$\mathbf{y}_{R,i}(n) = \mathbf{h}_i x_S(n) + \mathbf{G}_{\bar{i}} \mathbf{w}_{T,\bar{i}} x_R(n) + \mathbf{n}_{R,i}(n),\ i = 1,2, \tag{1}$$

$$\mathbf{y}_k(n) = \mathbf{g}_{i,k}^T \mathbf{w}_{T,i} x_R(n) + n_k(n), k = 1,2, \tag{2}$$

where $x_S(n) = \sum_{k=1}^2 x_k(n)$ is the composite AP message symbol with $\mathbb{E}\left[|x_S(n)|^2\right] = P_s$ when P_s is the source transmit power. The message $x_k(n)$ is destined to the k-th user with $\mathbb{E}\left[|x_k(n)|^2\right] = \rho_k P_s$ and $\sum_{k=1}^2 \rho_k = 1$. The cooperative HDRs apply the received beamforming to the signal in (2) as $\mathbf{w}_{R,i}^H \mathbf{y}_{R,i}(n)$, decode two messages with SIC manner from $x_1(n)$ to $x_2(n)$, re-encode these message signals $x_R(n+1) = \sum_{k=1}^2 x_k(n)\sqrt{1/P_s}$ and transmits $x_R(n+1)$ with beamforming as $\mathbf{w}_{T,i} x_R(n+1)$ at the time slot $n+1$. The vector $\mathbf{n}_{R,i}(n)$

denotes an $M \times 1$ additive noise vector at the receiver antennas of R_i with $\mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ distribution, while $n_k(n)$ denotes the additive noise at U_k antenna with $\mathcal{CN}(0,1)$ distribution.

The k-th user terminal uses the observation from the time slot n in (2) to decode the two messages in the SIC manner from $x_1(n)$ to $x_2(n)$. For the user k, the signal-to-interference-plus-noise ratio (SINR) for the signal $x_j(n),\ j \leq k$ through the i-th HDR is as follows:

$$\gamma_j^{k,i} = \frac{|\mathbf{g}_{i,k}\mathbf{w}_{T,i}|^2 \rho_j}{|\mathbf{g}_{i,k}\mathbf{w}_{T,i}|^2 \sum_{m=j+1}^{2} \rho_m + 1}. \tag{3}$$

On the other hand, each HDR is interfered by the other HDR as in (2), though they decode the messages with the same SIC manner as is carried out at user terminals. The SINR for $x_j(n)$ at the i-th HDR is given as follows

$$\gamma_j^{R,i} = \frac{|\mathbf{w}_{R,i}^H \mathbf{h}_i|^2 \rho_j P_s}{|\mathbf{w}_{R,i}^H \mathbf{G}_{\bar{i}} \mathbf{w}_{T,\bar{i}}|^2 + |\mathbf{w}_{R,i}^H \mathbf{h}_i|^2 \sum_{m=j+1}^{2} \rho_m P_s + 1}. \tag{4}$$

For the j-th message of time slot n $x_j(n)$, it should be decoded at the HDR and at both of the user terminals. Therfore, R_j^i, the transmission rate of $x_j(n)$ through the i-th HDR is determined by $\log(1 + \min[\gamma_j^{R,i}, \gamma_j^{1,i}, \gamma_j^{2,i}])$ with the DF protocol considered. Then, the average throughput from the two time slots through the HDRs is

$$\begin{aligned} R &= \frac{\sum_{i=1,j=1}^{2} R_j^i}{2} \\ &= \frac{1}{2} \sum_{i=1,j=1}^{2} \log(1 + \min[\gamma_j^{R,i}, \gamma_j^{1,i}, \gamma_j^{2,i}]). \end{aligned} \tag{5}$$

3. Beamformer Design

We design the beamformer vector set $\mathbf{w}_{R,i}$ and $\mathbf{w}_{T,i}$ for $i = 1,2$ such that the average throughput R is maximized with the two relays and two NOMA user terminals decoding the signals with the SIC manner. However, the fact that the relay R_i transmit beamformer vector $\mathbf{w}_{T,i}$ affects the SINRs $\gamma_j^{R,\bar{i}}, j = 1,2$ at $R_{\bar{i}}$ makes the optimization problem complicated and does not yield a straight-forward solution. Depending on the way they handle the cross-relay-interference (CRI), we propose two sub-optimal beamformer designs in the following subsections. Both designs let the relay receive beamformers $\mathbf{w}_{R,i}, i = 1,2$ to zero force the CRI perfectly and leave the transmit beamformers $\mathbf{w}_{T,i}, i = 1,2$ to take care of the NOMA user SINRs. This condition requires $\mathbf{w}_{R,i} = (\frac{\mathrm{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)}{\|\mathrm{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)\|})^H, i = 1,2$ to make the SINR at the relays in (4) to be free of CRI as

$$\gamma_j^{R,i} = \frac{|\mathbf{w}_{R,i}^H \mathbf{h}_i|^2 \rho_j P_s}{|\mathbf{w}_{R,i}^H \mathbf{h}_i|^2 \sum_{m=j+1}^{2} \rho_m P_s + 1}. \tag{6}$$

3.1. Multi-Dimensional Search Design with CRI Zero Forcing

Then, a relay transmit beamformer impacts the second hop link SINRs directly as in (3), while it indirectly affects the SINR of the other relay through $\mathbf{w}_{R,i}$ as in (6). This implies that the two pairs of receive and transmit beamformers should be designed jointly. The second hop design principle of [29] can be generalized to the beamforming of NOMA networks with a multiple antenna AP, where the beamformed channel gains of NOMA users should be equalized for the maximum NOMA sum rate to be achieved. Here, the transmit beamformer of a relay has an additional requirement of improving the SINR of the other relay. With the above considerations at hand, we may approach the problem with the following steps. Suppose the two receive beamformers $\mathbf{w}_{R,i}, i = 1,2$ are defined and we want to optimize the two transmit beamformers $\mathbf{w}_{T,i}, i = 1,2$. For $i = 1,2$ and $j = 1,2$, we may define matrices $\bar{\mathbf{G}}_{i,j} = \mathbf{g}_{i,j}\mathbf{g}_{i,j}^H$, $\bar{\mathbf{G}}_i = \mathbf{G}_{\bar{i}}^H \mathbf{w}_{R,i}\mathbf{w}_{R,i}^H \mathbf{G}_{\bar{i}}$ and $\mathbf{W}_i = \mathbf{w}_{T,i}\mathbf{w}_{T,i}^H$ to form the following two SDR problems with non negative definite matrices $\mathbf{W}_i, i = 1,2$. The connection between metrics in (3,6) and the formation of SDR problem can be found in references such as [13,30]. The simplified algorithm in Section 3.2 avoids the SDR formation and the subsequent alternation to narrow down the threshold ranges.

$$\begin{aligned} \min_{\mathbf{W}_1 \in \mathbb{C}^{M\times M}} \quad & \mathbf{Tr}[\mathbf{W}_1] \\ \text{s.t.} \quad & \mathbf{Tr}[\bar{\mathbf{G}}_{1,j}\mathbf{W}_1] \geq \gamma_1^*,\ j = 1,2, \\ & \mathbf{Tr}[\bar{\mathbf{G}}_1\mathbf{W}_1] \leq \delta_1, \\ & \operatorname{rank}\{\mathbf{W}_1\} = 1. \end{aligned} \tag{7}$$

$$\begin{aligned} \min_{\mathbf{W}_2 \in \mathbb{C}^{M\times M}} \quad & \mathbf{Tr}[\mathbf{W}_2] \\ \text{s.t.} \quad & \mathbf{Tr}[\bar{\mathbf{G}}_{2,j}\mathbf{W}_2] \geq \gamma_2^*,\ j = 1,2, \\ & \mathbf{Tr}[\bar{\mathbf{G}}_2\mathbf{W}_2] \leq \delta_2. \\ & \operatorname{rank}\{\mathbf{W}_2\} = 1. \end{aligned} \tag{8}$$

Here, γ_i^* is the threshold applied to all UT channels from the R_i uniformly to control the end throughput R_1^i and R_2^i. The CRI at R_i is controlled by $\delta_{\bar{i}}$ to affect the SINR $\gamma_j^{R,i}, j = 1,2$. Overall, two pairs of SDR in (7) and (8) pursue the balance points between the two UT SINRs and the CRI toward the other one in the relay pair. Reducing δ_i puts a restriction on achieving optimal γ_i^*, while it improves the CRI level at $R_{\bar{i}}$ and results in enhanced $\gamma_j^{R,\bar{i}}$ for $j = 1,2$. On the other hand, increasing δ_i works in the opposite directions and that it allows a chance to make a trade-off. Note that the power of $\|\mathbf{w}_{T,i}\|^2, i = 1,2$ can be compared with the allowed relay power P_r to test if the given set of values γ_i^* and δ_i can be further tightened as shown in Algorithm 1.

Now, all the above discussions can be integrated into the following twin optimizations as summarized in Algorithm 1, where the two pairs of beamformer vectors for two relays are found in an iterative fashion. At first, the relay receive beamformer vectors $\mathbf{w}_{R,i}, i = 1,2$ are set to the directions of the first hop channels $\mathbf{h}_i, i = 1,2$, while they are orthogonally projected at the end of the while loop to null out the CRI for the given-pair relay transmit vector $\mathbf{w}_{T,i}, i = 1,2$ set. In the while loop of the algorithm, two two-dimensional (2-D) searches are performed concurrently to find the pairs $\mathbf{w}_{T,i}, i = 1,2$, where the two 2-D tiles, each formed by δ_i and γ_i^*, are shrunk down in every iteration until the total throughput of the SR-NOMA is saturated. In each of the loop implementations, the mid points of the two 2-D areas are tested with the problems (7) and (8), where the magnitude squares of the resulting vectors $\mathbf{w}_{T,i}, i = 1,2$ are compared to the relay transmit power P_r to determine the next round of 2-D areas with each dimension shrink. The parameter ϱ $(0 \leq \varrho \leq 1)$ controls the amount that the 2-D areas shrink by after each iteration, where the shrinkage rate decelerates as the parameter approaches one. Note that two 2-D tiles are not obliged to converge to the same area, though the twin algorithms work in the same fashion. As the twin algorithms progress, the two relay transmit vectors $\|\mathbf{w}_{T,i}\|^2, i = 1,2$ approach P_r since the pair δ_i and γ_i^* are tightened in every while loop. Finally, note that the computational complexity of the twin iterative algorithm relies heavily on the SDR and the associated 2D search, which leads us to consider the one-shot algorithm in the following Section 3.2.

Algorithm 1: The twin iterative 2-D search algorithms for the SR-CNOMA

for $i = 1,2$ **do**
  Set $\gamma_{i,min} = 0$, $\gamma_{i,max} = \min_{j=1,2} P_r\|\mathbf{g}_{i,j}\|^2$ and $\delta_{i,min} = 0$, $\delta_{i,max} = P_r\sigma_0^2(\mathbf{G}_{\bar{i}})$.
  Set $\mathbf{w}_{R,i} = \mathbf{h}_i^* / \sqrt{\|\mathbf{h}_i\|}$.
end for
while ($\gamma_{i,max} - \gamma_{i,min}$ and $\delta_{i,max} - \delta_{i,min}$ for $i = 1,2$ do not reduce more than a set of marginal values) **do**
  for $i = 1,2$ **do**
    Set $\gamma_i^* = (\gamma_{i,min} + \gamma_{i,max})/2$ and $\delta_i = (\delta_{i,min} + \delta_{i,max})/2$.
  end for
  Perform the optimizations in (7) and (8).
  for $i = 1,2$ **do**
    if ($\|\mathbf{w}_{T,i}\|^2 \le P_r$) **then**
      Set $\gamma_{i,min} = (\varrho\gamma_{i,min} + (1-\varrho)\gamma_i^*)$ and $\delta_{i,max} = ((1-\varrho)\delta_{i,max} + \varrho\delta_i)$.
    else
      Set $\gamma_{i,max} = ((1-\varrho)\gamma_{i,max} + \varrho\gamma_i^*)$ and $\delta_{i,min} = (\varrho\delta_{i,min} + (1-\varrho)\delta_i)$.
    end if
  end for
  for $i = 1,2$ **do**
    Set $\mathbf{w}_{R,i} = (\frac{\mathsf{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)}{\|\mathsf{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)\|})^H$.
  end for
end while
for $i = 1,2$ **do**
  Set $\gamma_i^* = \gamma_{i,min}$ and $\delta_i = \delta_{i,max}$.
end for
Perform the optimizations in (7) and (8).
for $i = 1,2$ **do**
  Set $\mathbf{w}_{R,i} = (\frac{\mathsf{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)}{\|\mathsf{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)\|})^H$.
  Scale $\mathbf{w}_{T,i}$ such that $\|\mathbf{w}_{T,i}\|^2 = P_r$.
  Calculate $\gamma_1^{1,i}, \gamma_1^{2,i}, \gamma_2^{1,i}, \gamma_2^{2,i}, \gamma_1^{R,i}$ and $\gamma_2^{R,i}$ using (3) and (6).
end for
Calculate $R = \frac{\sum_{i=1,j=1}^{2} R_j^i}{2}$ using (5).

3.2. One-Shot Design with CRI Zero Forcing

Algorithm 1 has iterative applications of SDR and associated multi-dimensional search, which may require quite a bit of computational load. Here, we propose a CRI zero-forcing base reference scheme implemented in one shot approach with much reduced computation. Algorithm 2 sets the transmit beamformer $\mathbf{w}_{T,i}$ of the i-th HDR to the direction of the channel vector $\mathbf{g}_{i,i}$ and the receive beamformer to the zero-forcing direction of CRI. The beamformer set computation of two HDRs is coupled only through CRI components. This formation considers the second hops (relays to NOMA user channels) as if only a single user is served so that the application of SDR is avoided and lets only the relay receive beamformers to take care of the CRI.

Algorithm 2: : The one-shot algorithm

1. **for** $i = 1, 2$ **do**
2. Set $\mathbf{w}_{T,i} = (\frac{\mathbf{g}_{i,i}}{\|\mathbf{g}_{i,i}\|})^H$.
3. Set $\mathbf{w}_{R,i} = (\frac{\mathrm{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)}{\|\mathrm{P}_{\mathbf{G}_{\bar{i}}\mathbf{w}_{T,\bar{i}}}(\mathbf{h}_i)\|})^H$.
4. Scale $\mathbf{w}_{T,i}$ such that $\|\mathbf{w}_{T,i}\|^2 = P_r$.
5. Calculate $\gamma_1^{1,i}, \gamma_1^{2,i}, \gamma_2^{1,i}, \gamma_2^{2,i}, \gamma_1^{R,i}$ and $\gamma_2^{R,i}$ using (3) and (4).
6. **end for**
7. Calculate $R = \frac{\sum_{i=1,j=1}^{2} R_j^i}{2}$ using (5).

4. Numerical Results

In this section, we present numerical results based on the average sum rate performance of the proposed optimization schemes and of the HDR based cooperative NOMA system in various parameter settings. The system model in Figure 1 is considered with the assumed statistical properties for the channel gains, where two relaying UTs and two NOMA users are served by an AP. For each curve in the simulations, ten thousand channels are generated independently to average out the sum rate over these channels. Furthermore, the noise power is normalized to one, and the signal power values are all relative to this noise power. The HDR-based scheme does the role of a baseline scheme, which requires only one relaying UT and two time slots for the cooperative NOMA transmission, and thus, it suffers no CRI, while it requires twice the spectral resources. Furthermore, it requires one SDR operation for the UT transmit beamforming for the two NOMA users.

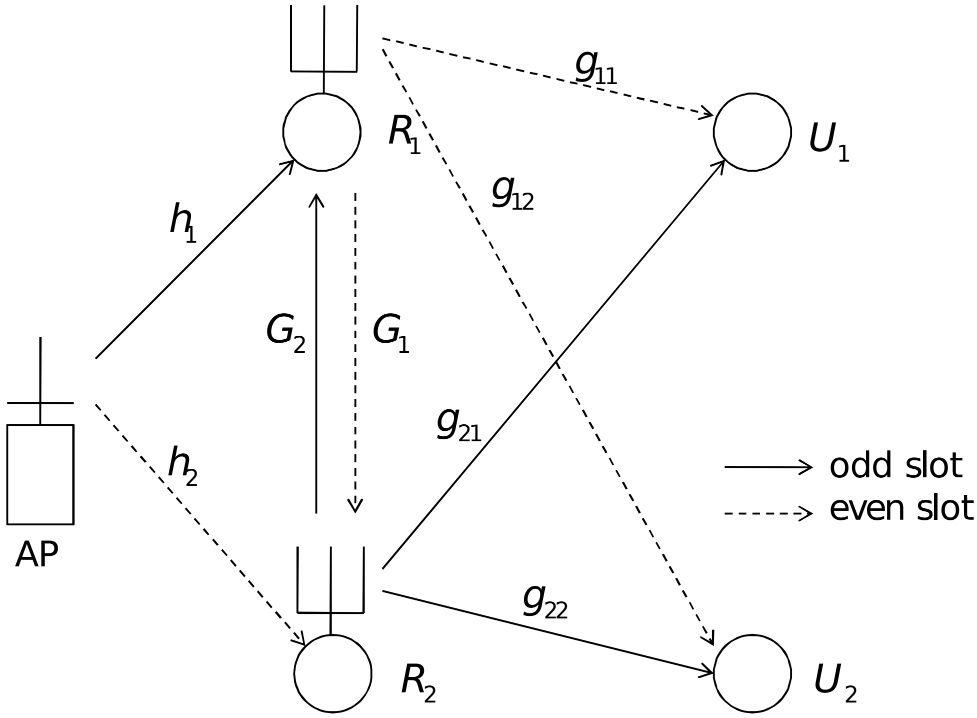

Figure 1. Cooperative non-orthogonal multiple-access transmission system through successive relaying.

In Figure 2, the average sum throughput (bit per second per hertz; BPS) of the various schemes described in this paper are displayed over the relaying UT transmit power (P_r) with three different AP transmit power (P_s) cases. The number of antennas at the relaying UTs is set to three ($M = 3$). This experiment shows how effectively the proposed schemes handle the CRI (as it depends on the relay transmit power P_r) and achieve better average throughput since the proposed schemes beat the HDR cases in all three cases and over the whole relay power range (P_r). As expected proposed two algorithms which exhibit their capability in handling CRI in all the three cases, while Algorithm 1 shows the extra benefit of NOMA beamforming gain obtained from the iterative applications of SDR in the algorithm, which is reflected in the additional throughput gain of Algorithm 1 over Algorithm 2. This gain of Algorithm 1 appears in the low-to-mid range of P_r, though it diminishes as P_r increases further and dominates the performance. The two proposed SR schemes successfully mimic the full duplex operation by suppressing the CRI very well and surpassing the average sum throughput of HDR cases. Different from the case of Figure 2, where the AP power P_s is fixed and the relaying UT power P_r varies, Figure 3 tests the case where P_s varies while P_r is fixed for three different values. Again, the two proposed schemes take the advantage of the AP transmit power into the performance gain so that they excel the HDR in performance. The amount of advantage of two Algorithms diminishes as P_r reduces, and Algorithm 2 suffers more than Algorithm 1 in this respect.

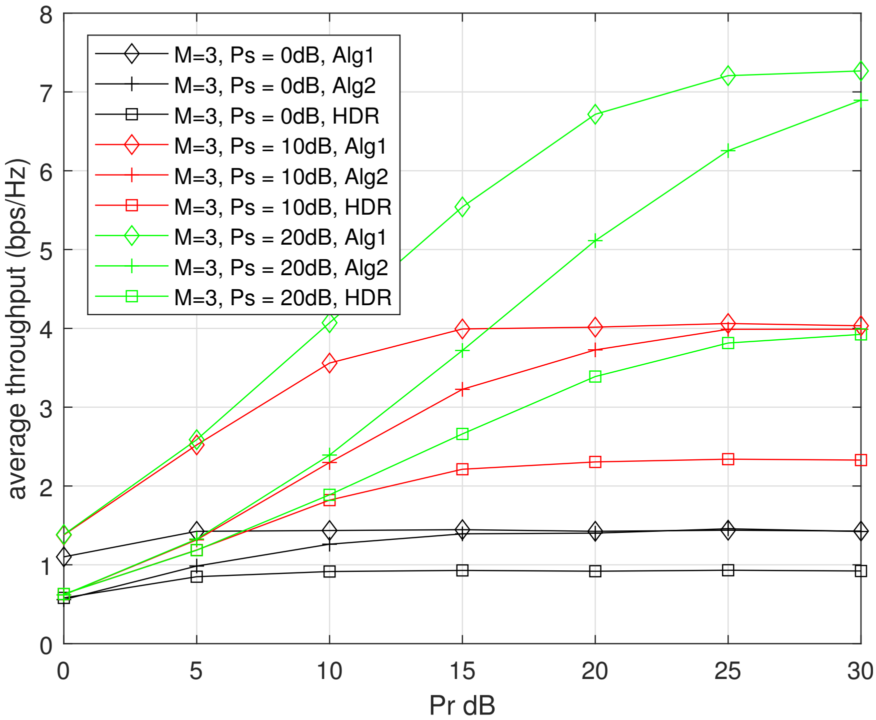

Figure 2. The average sum throughput in bits per second per hertz (BPS) comparison over the relaying UT transmit power with different AP transmit powers (P_s). Here, $M = 3$.

Now in Figure 4, the average sum throughput of proposed schemes over the number of relaying UT transmit antennas M are compared with three different sets of fixed P_s and P_r values. First, consider the case of $P_s = 0$ (dB) and $P_r = 20$ (dB) in black curves, which reminds us of the cases of Figure 2, where the advantage of Algorithm 1 over Algorithm 2 disappears as P_r becomes excessive and the suppression of CRI is challenging. Second, the case is reversed in red curves and that we set $P_s = 20$ (dB) and $P_r = 0$ (dB). The interpretation of this case needs the observations from Figure 3 as the performance of Algorithm 2 suffers in low P_r values. Here, Algorithm 2 performs even worse than the HDR case as M grows into large values, while the gain of Algorithm 1 is limited as well. Finally, the most interesting case of green curves corresponds to the setting of $P_s = 10$ (dB) and $P_r = 10$ (dB). Here, the curves of Algorithm 2 flatten and become degrading after M increases by more than 6. This weird event can be explained by the channel hardening property of the multiple antenna channel at the massive value of M. In this regime with a massive number of antennas, the cross-UT channels ($\mathbf{G}_1$ and $\mathbf{G}_2$ in Figure 1) become almost orthogonal to the AP to relaying UT channels ($\mathbf{h}_2$ and $\mathbf{h}_1$ in Figure 1), and thus, the zero forcing operations in Algorithm 1 and Algorithm 2 do not help much. Instead, the UT transmit beamforming of

Algorithm 2 becomes ineffective as they simply point to the one of target UTs, respectively. This handicap of Algorithm 2 is reflected as the degradation of its performance at the large M regime.

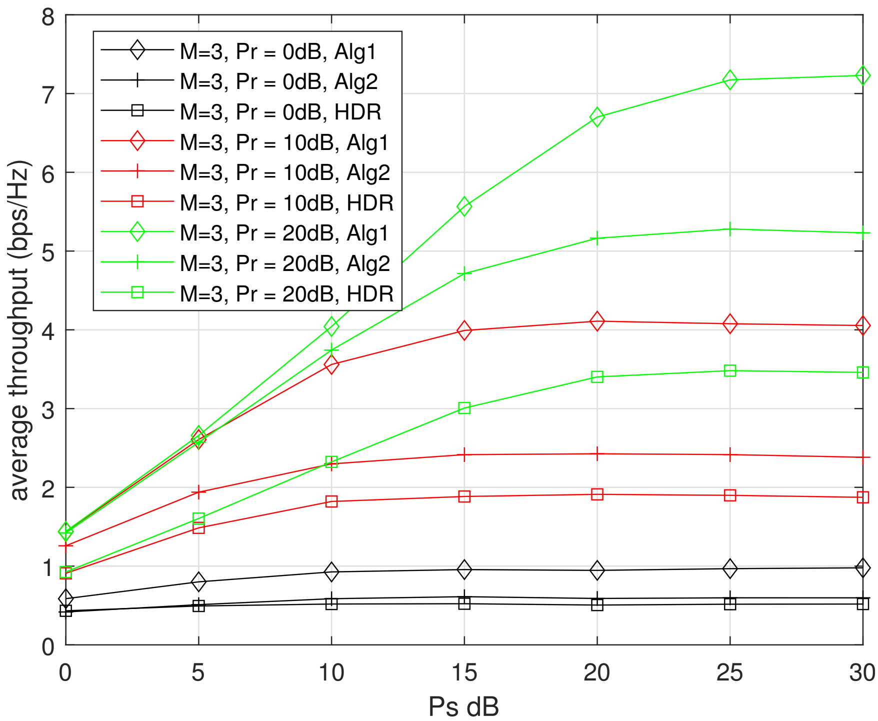

Figure 3. The average sum throughput in bits per second per hertz (BPS) comparison over the AP transmit power with different relaying UT transmit powers (P_r). Here, $M = 3$.

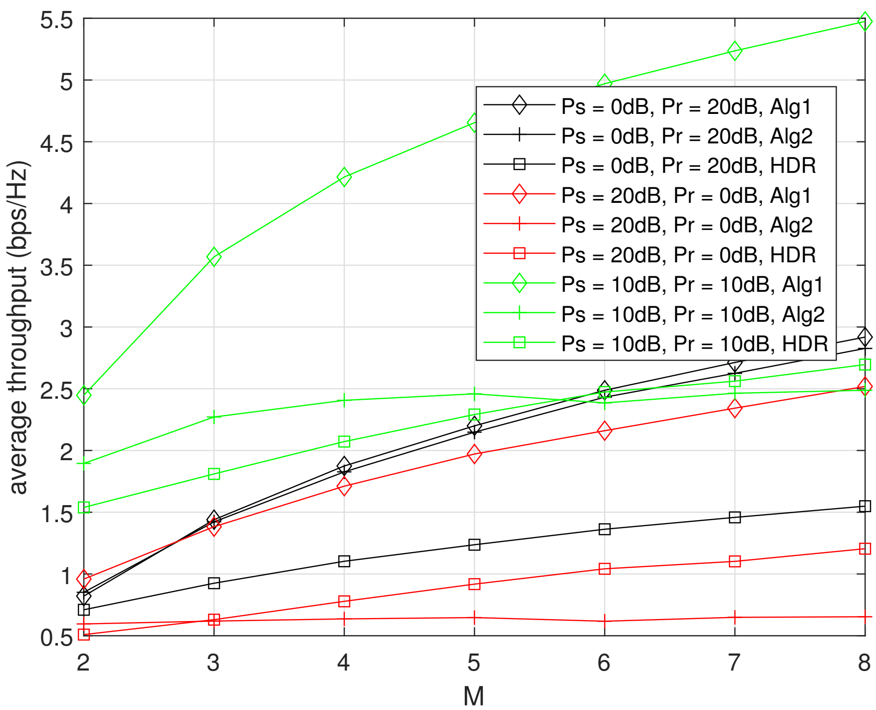

Figure 4. The average sum throughput in bits per second per hertz (BPS) comparison over the number of UT antenna set with different AP transmit powers (P_s) and different relaying UT transmit powers (P_r).

5. Conclusions

Successive relaying-based cooperative non orthogonal multiple access (CNOMA) transmission schemes are proposed. Two half-duplex relaying user terminals deliver the AP signals toward two NOMA UTs one-by-one. Thus, they emulate the full-duplex relaying transmission and the end-to-end transmission of one NOMA data block is carried out in a single time slot. The removal of cross-relay interference between the two HD relays is the key technical challenge and the beamforming with multiple antennas at the relaying UTs is employed for this purpose. Furthermore, the relaying UT beamforming does the additional role of maximizing the CNOMA transmission efficiency. Among the two CRI zero forcing beamforming schemes proposed, the first method employs semi-definite relaxation (SDR) with a four dimensional search, where the four dimensions reflect the four parameters of the CRI levels at the relaying UTs and signal strengths at the NOMA UTs. This scheme alternately updates beamformer sets until the 4-D search range is reduced and the sum throughput rate of CNOMA saturates. The second method avoids the SDR and alternation and that the beamformer set is found in one shot calculation. The performances of these two schemes and HDR relay-based cooperative NOMA scheme are compared in the numerical experiments. The two proposed schemes outperform the HDR-based cooperative NOMA in most of the cases though the advantage of the first scheme over the second scheme diminishes in the high relaying power regime (high CRI regime), where the two schemes perform almost the same. When the number of antennas increases into the range of massive ones, the second scheme fails to outperform that of the HDR-based scheme since it lacks the optimization of CNOMA transmission of the second hop. The results of this work indicate that the coverage area of the NOMA service can be extended by a set of beamforming schemes for the cooperative relaying UTs with no extra spectral resources.

Author Contributions: D.H. proposed the beamforming principles and associated algorithms for the SR cooperative NOMA wireless networks; J.Y. and S.S.N. supplemented in improving the algorithms and derived and analyzed all simulation results. Furthermore, H.-K.S. provided the experimental materials for better computational simulations and revised critical errors in the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (2020R1A6A1A0 3038540) and the Institute of Information and communications Technology Planning and Evaluation (IITP) under the metaverse support program to nurture the best talents (IITP-2023-RS-2023-00254529) grant funded by the Korean government (MSIT).

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Saito, Y.; Kishyama, Y.; Benjebbour, A.; Nakamura, T.; Li, A.; Higuchi, K. Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access. In Proceedings of the 2013 IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, Germany, 2–5 June 2013; pp. 1–5.
2. Wang, Q.; Zhang, R.; Yang, L.-L.; Hanzo, L. Non Orthogonal Multiple Access: A Unified Perspective. *IEEE Wirel. Commun. Mag.* **2018**, *25*, 10–16. [CrossRef]
3. Liu, Y.; Pan, C.; Nallanathan, A.; Elkashan, M.; Hanzo, L. Multiple-Antenna-Assisted Non-Orthogonal Multiple Access. *IEEE Wirel. Commun. Mag.* **2018**, *25*, 17–23. [CrossRef]
4. Huang, Y.; Zhang, C.; Wang, J.; Jing, Y.; Yang, L.; You, Z. Signal Processing for MIMO-NOMA: Present and Future Challenges. *IEEE Wirel. Commun. Mag.* **2018**, *25*, 32–38. [CrossRef]
5. Fang, F.; Zhang, H.; Cheng, J.; Leung, V.C.M. Energy-Efficient Resource Allocation for Downlink Non-Orthogonal Multiple Access Network. *IEEE Trans. Commun.* **2016**, *64*, 3722–3732. [CrossRef]
6. Bao, W.; Chen, H.; Li, Y.; Vucetic, B. Joint Rate Control and Power Allocation for Non-Orthogonal Multiple Access Systems. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 2798–2811. [CrossRef]
7. Zhu, J.; Wang, J.; Huang, Y.; He, S.; You, X.; Yang, L. On Optimal Power Allocation for Downlink Non-Orthogonal Multiple Access Systems. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 2744–2757. [CrossRef]

8. Pliatsios, D.; Sarigiannidis, P. Resource allocation combining heuristic matching and particle swarm optimization approaches: The case of downlink non-orthogonal multiple access. *Information* **2019**, *10*, 336. [CrossRef]
9. Zhao, C.; Cai, Y.; Liu, A.; Zhao, M.; Hanzo, L. Mobile edge computing meets mmWave communications: Joint beamforming and resource allocation for system delay minimization. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 2382–2396. [CrossRef]
10. Zhu, Z.X.L.; Choi, J.; Xia, P.; Xia, X. Joint power allocation and beamforming for non-orthogonal multiple access (NOMA) in 5G millimeter wave communications. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 2961–2974. [CrossRef]
11. Zhu, L.; Zhang, J.; Xiao, Z.; Cao, X.; Wu, D.; Xia, X. Millimeter-wave NOMA with user grouping, power allocation and hybrid beamforming. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 5065–5079. [CrossRef]
12. Hanif, M.F.; Ding, Z.; Ratnarajah, T.; Karagiannidis, G.K. A Minorization-Maximization Method for Optimizing Sum rate in the Downlink of Non-Orthogonal Multiple Access Systems. *IEEE Trans. Veh. Technol.* **2016**, *64*, 76–88. [CrossRef]
13. Ding, Z.; Adachi, F.; Poor, H.V. The Application of MIMO to Non-Orthogonal Multiple Access. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 537–552. [CrossRef]
14. Ding, Z.; Schober, R.; Poor, H.V. A General MIMO framework for NOMA downlink and uplink transmission based on signal alignment. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 4438–4454. [CrossRef]
15. Chen, Z.Y.; Ding, Z.; Dai, X. Beamforming for Combating Inter-cluster and Intra-cluster Interference in Hybrid NOMA Systems. *IEEE Access* **2016**, *4*, 4452–4463. [CrossRef]
16. Jeong, Y.; Lee, C.; Kim, Y. Power minimizing beamforming and power allocation for MISO-NOMA systems. *IEEE Trans. Veh. Technol.* **2019**, *68*, 6187–6191. [CrossRef]
17. Sun, X.; Nan, Y.; Yan, S.; Ding, Z.; Ng, D.; Shen, C.; Zhong, Z. Joint beamforming and power allocation in downlink NOMA multiuser MIMO networks. *IEEE Trans. Wirel. Commun.* **2018**, *17*, 5367–5381. [CrossRef]
18. Al-Obiedollah, H.; Cumanan, K.; Thiyagalingam, J.; Burr, A.G.; Ding, Z.; Dobre, O.A. Energy efficient beamforming design for MISO non-orthogonal multiple access systems. *IEEE Trans. Commun.* **2019**, *67*, 4117–4131. [CrossRef]
19. Alavi, F.; Cumanan, K.; Ding, Z.; Burr, A.G. Beamforming techniques for nonorthogonal multiple access in 5G cellular networks. *IEEE Trans. Veh. Technol.* **2018**, *67*, 9474–9487. [CrossRef]
20. Nandan, N.; Sudhan, M.; Wu, H. Secure beamforming for MIMO-NOMA-based cognitive radio network. *IEEE Commun. Lett.* **2018**, *8*, 1708–1711. [CrossRef]
21. Feng, Y.; Yan, S.; Yang, Z.; Yang, N.; Yuan, J. Beamforming Design and Power Allocation for Secure Transmission with NOMA. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 2639–2651. [CrossRef]
22. Xu, Y.; Shen, C.; Ding, Z.; Sun, X.; Yan, S.; Zhu, G.; Zhong, Z. Joim Beamforming and Power-Spliting Control in Downlink Cooperative SWIPT NOMA Systems. *IEEE Trans. Signal Process.* **2017**, *65*, 4874–4886. [CrossRef]
23. Alsaba, Y.; Leow, C.Y.; Rahim, S.K.A. Full-Duplex Cooperative Non-orthogonal multiple access with beamforming and energy harvesting. *IEEE Access* **2018**, *6*, 19726–19738. [CrossRef]
24. Zhang, Z.; Ma, Z.; Xiao, M.; Ding, Z.; Fan, P. Full-Duplex Device-to-device-aided Cooperative Non-orthogonal multiple access. *IEEE Trans. Veh. Technol.* **2017**, *66*, 4467–4471. [CrossRef]
25. Kim, J.; Lee, I. Capacity Analysis of Cooperative relaying Systems Using Non-Orthogonal Multiple Access. *IEEE Commun. Lett.* **2015**, *19*, 1949–1952. [CrossRef]
26. Ding, Z.; Peng, M.; Poor, H.V. Cooperative Non-Orthogonal Multiple Access in 5G Systems. *IEEE Commun. Lett.* **2015**, *19*, 1462–1465. [CrossRef]
27. Yang, Z.; Ding, Z.; Wu, Y.; Fan, P. Novel Relay Selection Strategies for Cooperative NOMA. *IEEE Trans. Veh. Technol.* **2017**, *66*, 10114–10123. [CrossRef]
28. Liau, Q.Y.; Leow, C.Y.; Ding, Z. Amplify-and-Forward Virtual Full-Duplex Relaying-Based Cooperative NOMA. *IEEE Wirel. Commun. Lett.* **2018**, *7*, 464–467. [CrossRef]
29. Hwang, D.; Yang, J.; Nam, S.S.; Jeong, J.; Song, H. Cooperative Non-Orthogonal Multiple Access Transmission through Full-Duplex and Half-Duplex Relays. *IEEE Wirel. Commun. Lett.* **2023**, *12*, 351–355. [CrossRef]
30. Sidiropoulos, N.D.; Davidson, T.N.; Luo, Z.Q. Transmit Beamforming for Physical-Layer Multicasting. *IEEE Trans. Signal Process.* **2006**, *54*, 2239–2251. [CrossRef]