Down-link NOMA with Successive Refinement for Binary Symmetric Source Transmission

Meng Cheng, Wensheng Lin, *Member, IEEE,* Tad Matsumoto, *Fellow, IEEE*

***Abstract*—This paper focuses on a lossy transmission of binary symmetric source (BSS) with down-link non-orthogonal multiple access (DL-NOMA). The transmitted binary sequences are lossy-compressed descriptions of the same BSS, which are overlapped in a specified signal format, corresponding to the receivers with pre-determined distortion requirements. Despite a higher spectrum efficiency achieved by DL-NOMA, redundancy may still remain if the overlapped descriptions are correlated. In this work, a system combining DL-NOMA and successive refinement, referred to as DN-SR, is proposed. Specifically, instead of re-constructing the source by a single description, the receiver can achieve a low distortion by first recovering the basic description, and then refining it with the help of another refinement description. If the two descriptions are made independent, transmission efficiency of the conventional DL-NOMA can be further improved. As the main contribution, this paper derives the outage probability of DN-SR in closed-form, assuming block Rayleigh fading channels. The advantage of DN-SR is numerically studied in terms of the system outage probability, compared to both the conventional DL-NOMA, and a modified DL-NOMA which exploits the correlation between descriptions, referred to as DN-CE. Finally, the optimal power allocation to the two descriptions is investigated for DN-SR, aiming to minimize the system outage probability.**

***Index Terms*—Down-link NOMA, successive refinement, lossy transmission, outage probability, power allocation.**

I. Introduction

Efficient point-to-multipoint (P2MP) communication plays an important role in wireless down-link systems, such as the evolved multimedia broadcast multicast service (eMBMS) specified by 3GPP [1], and the advanced television system committee (ATSC) 3.0 standard [2], [3]. Due to the ambitious goal towards connecting massive devices in future wireless networks, such as 5th generation (5G) and beyond 5G (B5G) systems, co-existence of the receivers having different detection capabilities and/or quality-of-service (QoS) requirements should be considered. Therefore, a lossy transmission model is focused on in this paper, where multiple descriptions with different distortion levels of the same source are transmitted via the down-link. As a conventional way of improving the spectrum efficiency, non-orthogonal transmission is applied which accommodates multiple communication pipes in the system sharing the same transmission resources [4], [5]. For exmaple in [6], digital TV signals having 4K and 8K resolutions are simultaneously broadcast with layer division multiplexing (LDM), known as one of the down-link non-orthogonal multiple access (DL-NOMA) schemes. It has been shown that expansion of the rate region is achieved with successive interference cancellation (SIC) [7].

Despite the benefits brought by DL-NOMA, the transmission efficiency may still be less than the optimal case if correlation remains among multiple descriptions at the transmitter. Considering a two-receiver case with SIC, the receiver near to the transmitter has to first detect the undesired signal and then eliminate its interference, before detecting its own signal [8]. In this case, such interference component detected at the near receiver will be always discarded after SIC. To further improve the transmission efficiency, this paper proposes a combination of DL-NOMA with successive refinement technique, referred to as DN-SR. Specifically, instead of broadcasting two descriptions with one-to-one correspondence to the receivers, the original source is first lossy-compressed into a basic and a refinement description. They can be made mutually independent with specified lossy compression schemes.[1] The basic description is detected by both receivers for re-constructing the source at a basic QoS, and the refinement description will be further utilized for the receiver aiming at a higher QoS. In this sense, the use of the basic description at the high-QoS receiver can alleviate the burden of transmitting the second description with a high rate [9].

The pioneer work combining non-orthogonal transmission and successive refinement can be found in [10], where Gaussian sources are transmitted over block Rayleigh fading channels. The optimal power allocation ratio, yielding the minimum average distortion, is calculated by assuming an equal rate for all descriptions. In [11], multiple descriptions of a Gaussian source are broadcast in both fast and block Rayleigh fading channels, and the optimal power allocation ratio is identified, given a linear distortion cost function. More discussions of the techniques in this category can be found in [12], where the effects such as bandwidth efficiency, operating signal-to-noise ratio (SNR) and diversity order are analyzed. The optimal power allocation ratio is also calculated by minimizing a convex distortion cost function. To the best of our knowledge, the previous work only investigate the Gaussian source, which are not applicable to binary source

M. Cheng is with the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, 923-1292, Japan (e-mail: m-cheng@jaist.ac.jp).

W. Lin is with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, Shaanxi, 710072, China (e-mail: linwest@nwpu.edu.cn).

T. Matsumoto is with with the School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), Nomi, Ishikawa, 923-1292, Japan, and the Center for Wireless Communications, University of Oulu, 90014 Oulu, Finland (Currently this position is frozen) (e-mail: matumoto@jaist.ac.jp, tadashi.matsumoto@ee.oulu.fi).

[1]However, practical code design for lossy compression is out of the scope of this paper.

transmission in practice.

For the sake of rate region expasion, a broadcast relay channel (BRC) model is studied in [13], where two descriptions are constructed at relay and forwarded to receivers having different QoS requirements, respectively. Similarly, it has been proven in [14] that broadcasting dedicated descriptions strictly achieves better performances than those with only one common description. However, only multiple description coding (MDC) is assumed in [13], [14], and the system performance in fading channels is not evaluated.

Motivated by the insufficiencies described above, this paper focuses on a non-orthogonal lossy transmission of binary symmetric source (BSS), where distortion is defined by the Hamming distortion measure [15]. With BSS, the instantaneous channel capacity is constrained by specific modulations, known as the constellation constrained capacity (CCC) [16], [17]. This assumption is more reasonable when evaluating practical systems, where transmit symbols are from a finite alphabet set. With block Rayleigh fading, the channel state over one-block duration is assumed to be constant. Therefore, instead of averaging the channel capacity, the system performance is evaluated by outage probabilities.

As in [18], [19], [20], the outage probability is calculated based on the rate region analysis, where Shannon's source-channel separation theorem is used. It is well known that the separation theorem holds for point-to-point (P2P) transmissions [15], but does not always hold for networks. In this case, results obtained by a separate source-channel coding can be regarded as a performance upper bound, as in [21]. However, according to [22], [23], the separation theorem still holds for some special networks under two conditions: 1) multiple descriptions are memory-less and mutually independent, and 2) each description is needed only by one dedicated receiver, or by multiple receivers with the same distortion requirement. In the proposed DN-SR system, the basic description is needed by both receivers with a same distortion requirement. Furthermore, the basic description is independent of the refinement description which is needed only by the high-QoS receiver. Therefore, the separation theorem holds for the proposed DN-SR system.

This paper also provides evaluation results of two counterpart schemes. First of all, the conventional DL-NOMA is analyzed, where the correlation between two descriptions are kept, as in [2], [3]. Source re-construction at each receiver only relies on the detected description dedicating to this receiver. Furthermore, in order to achieve a better decoding performance, a modified DL-NOMA is investigated such that the correlation knowledge is exploited at the receiver, which is referred to as DN-CE. Since two descriptions sent by the conventional DL-NOMA and DN-CE are correlated, Shannon's separation theorem may not hold. Results obtained in such cases are regarded as the outage probability upper bounds [24]. For the transmission towards the high-QoS receiver, Slepian-Wolf bound [25] is analyzed, which defines the rate constraints for transmitting correlated sequences in a loss-less system. Some practical techniques for correlation exploitating can be found in [18], [26]. In [26], significant gains in terms of the required SNR can be observed with the proposed technique in both additive white Gaussian noise (AWGN) and Rayleigh fading channels, compared to those without exploiting the source correlation. The outage probability bound based on Slepian-Wolf theorem is first derived in [18] for the proposed relay network, under block Rayleigh fading. Although there is a 2 dB SNR gap between the theoretical bound and the practical result, the performance tendencies are consistent. Hence, it is reasonable to consider the technique shown in [26] to improve the performance of DL-NOMA, as in DN-CE. Note that the practical code design for DN-CE is out of the scope of this paper.

The main contributions of this paper are summarized as follows:

1) A closed-form expression of outage probability is derived for the proposed DN-SR system with BSS;
2) Comparative studies are provided with the conventional DL-NOMA and DN-CE, where the performance gain of DN-SR has been observed from numerical results;
3) The optimal power allocation ratio for DN-SR is calculated which minimizes the outage probability by formulating the problem with a convex optimization.

The organization of this paper is as follows. First of all, models of the proposed DN-SR and the comparative schemes are introduced in Section II, followed by the preliminaries of broadcast channel model and successive refinement. As the main body, rate region analyses and outage probability derivations are provided in Section III for the proposed DN-SR, as well as the reference schemes. Moreover, numerical evaluations for the outage probability are provided for those schemes in Section IV. In Section V, the optimal power allocation ratio for DN-SR is calculated by solving a convex optimization problem. Finally, this work is concluded in Section VI with some technical remarks.

II. System Model

A. Proposed DN-SR System

The proposed DN-SR system is described in this subsection. First of all, let SC_i, CC_i and Map_i denote the source encoder, the channel encoder and the mapper, respectively, with $i = \{1, 2\}$, and $(\cdot)^{-1}$ denoting the inverse operation of its argument in the diagram. Let X denote the binary variable from the original BSS to be transmitted, with its entropy $H(X) = 1$.

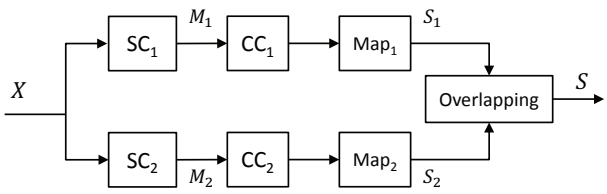


Fig. 1. Transmitter diagram of DN-SR.

According to the transmitter diagram shown in Fig. 1, two descriptions of X are generated via lossy compression, i.e., the basic description M_1 and the refinement description M_2, respectively, according to the successive refinement concept.

To be specific, M_1 includes complete information for re-constructing X with a distortion D_1, and such result can be further refined with the help of M_2 to achieve a distortion D_0, where $D_0 \leq D_1$. M_1 and M_2 are then channel-encoded and mapped in parallel, yielding the symbol sequences S_1 and S_2, respectively. The broadcast signal S is an overlapped version of S_1 and S_2 in power domain [27], given by

$$S = \alpha S_1 + \tilde{\alpha} S_2, \tag{1}$$

where $\alpha \in (0,1)$ is an adjustable power allocation ratio, and $\tilde{\alpha} = 1-\alpha$. Obviously, S_1 and S_2 should have the same length.

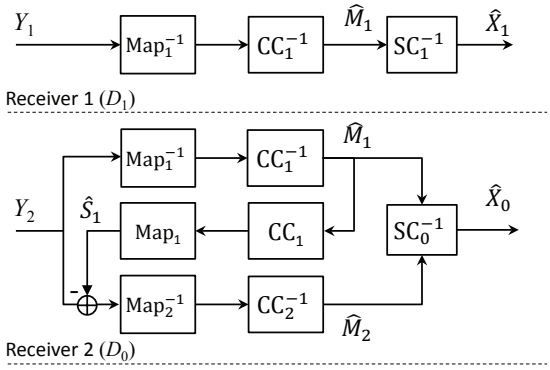


Fig. 2. Receiver diagram of DN-SR.

As shown in Fig. 2, two receivers, one located far from and another near to the transmitter, are deployed in the proposed system, denoted by Receiver 1 and Receiver 2, respectively. They are required to re-construct the original source with distortion D_1 and D_0. Wireless links from the transmitter to Receiver 1 and Receiver 2 are denoted by Link 1 and Link 2, respectively, each suffering from independent block Rayleigh fading. Therefore, the received signals at two receivers can be expressed by

$$\begin{aligned} Y_1 &= h_1 S + n_1, \\ Y_2 &= h_2 S + n_2, \end{aligned} \tag{2}$$

where h_1 and h_2 denote the complex channel gains with $\mathbb{E}\left[|h_1|^2\right] \leq \mathbb{E}\left[|h_2|^2\right]$, n_1 and n_2 are the two-dimensional AWGN noises following the distribution $\mathcal{N}(0,\sigma^2)$ in each dimension. At Receiver 1, only single user detection for S_1 is performed, where S_2 is simply regarded as the noise. Therefore, the transmission of Link 1 can be always regraded as P2P. The re-constructed data sequence $\hat{X}_1$ is only based on $\hat{M}_1$, which is the output of the channel decoder CC_1^{-1}. In contrast, S_1 and S_2 are both detected at Receiver 2 with SIC, and the re-constructed data sequence $\hat{X}_0$ is based on both $\hat{M}_1$ and $\hat{M}_2$, where $\hat{M}_2$ is the detected refinement description.

Let R_{si} denote the rate for transmitting M_i after lossy compression, and R_{ci} represent the normalized spectrum efficiency, i.e., the product of channel coding rate and modulation order. Therefore, the total rate of the i-th signal layer is given by $R_i = R_{si}R_{ci}$. Since the two signal layers share the same transmission resources, the symbol-level frame length should be always kept the same. Therefore, we define $\eta_i = R_{si}/R_{ci}$, and η_1 should be equal to η_2.

B. Schemes for Comparison

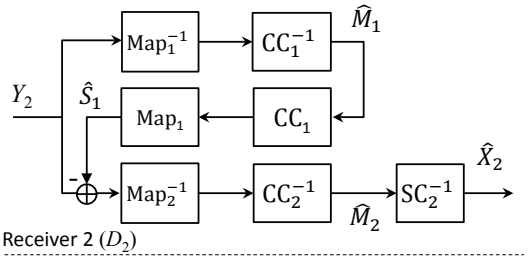


Fig. 3. Receiver diagram of conventional DL-NOMA.

Considering a same original source X, the conventional DL-NOMA is first investigated. It adopts a different source coding scheme from DN-SR, i.e., M_2 is no longer generated as a refinement description of M_1, but the dedicated data which Receiver 2 aims at, as shown in Fig. 3. In other words, source re-construction at Receiver 2 only relies on $\hat{M}_2$, while still keeping the same distortion requirement D_0 as in DN-SR for a fair comparison. Since D_0 used in DN-SR implies the distortion achieved by a joint decoding of M_1 and M_2, which is not the case for DL-NOMA, D_2 is alternatively used here to denote the required distortion at Receiver 2, with $D_2 = D_0$.

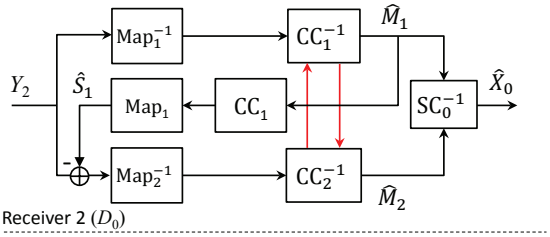


Fig. 4. Receiver diagram of DN-CE.

Since Receiver 2 only relies on $\hat{M}_2$ to achieve a low distortion in the conventional DL-NOMA, M_2 should be transmitted with a high rate. In such a case, M_2 is most likely correlated with M_1, because they are generated from the same source. To exploit such correlation for a better decoding performance, an improved scheme of DL-NOMA is investigated, referred to as DN-CE. According to Fig. 4, the correlation knowledge is exploited at Receiver 2 by exchanging information between the two component decoders [26], [28]. The source re-construction is finally performed based on $\hat{M}_1$ and $\hat{M}_2$, where the Slepian-wolf theorem is used for theoretical rate region analysis.

Note that making comparison between DN-CE and DN-SR is of great importance, because they are regarded as different approaches to enhance the conventional DL-NOMA. Despite different source coding methods, fair comparison is made possible by setting the same distortion requirements to the two schemes with the same original source, and the total transmission resources are kept equal. However, those constraints result in different spectrum efficiency and SNR requirement, as detailed in the following discussions.

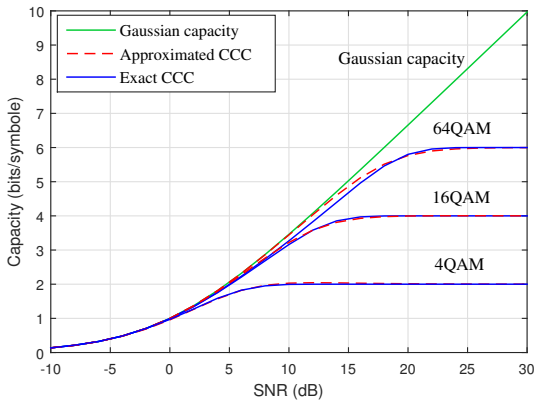


Fig. 5. Accuracy evaluation of exact and approximated CCC results for 4QAM, 16QAM and 64QAM.

C. Preliminaries

1) Broadcast Channel: The proposed transmission can be seen as a broadcast channel [9]. According to the previous assumption of near/far receivers in Section II-A, the average received SNR at Receiver 2 is larger, where SIC detection is performed. Let γ_1 and γ_2 denote the instantaneously received SNRs of the two links, the capacities of the Gaussian broadcast channel with SIC is given by

$$R_1 \leq C\left(\frac{\alpha\gamma_1}{\tilde{\alpha}\gamma_1+1}\right), \tag{3}$$

$$R_2 \leq C(\tilde{\alpha}\gamma_2), \tag{4}$$

where $C(\cdot)$ represents the capacity function of a P2P transmission. With a Gaussian codebook, $C(\gamma_i) = \log_2(1+\gamma_i)$. However, this paper assumes BSS with a specific $\mathcal{M}$-ary quadrature amplitude modulation (QAM). Therefore, the CCC rather than the Gaussian capacity is utilized for the sake of more accurate performance evaluations. Since that closed-form expression for calculating the explicit CCC is not available, an approximated equation derived in [29] is utilized to facilitate the outage derivation of our proposed system, which is given by

$$C(\gamma_i) = \log_2(1+\gamma_i) - \frac{1}{2}\log_2\left[1+\left(\frac{\gamma_i}{\mathcal{M}}\right)^2\right]. \tag{5}$$

Inversely, the required instantaneous SNR for achieving the capacity C can be calculated by $\gamma_i = Z(C)$, where

$$Z(C) = \frac{\mathcal{M}2^C\sqrt{\mathcal{M}^2-4^C+1}-\mathcal{M}^2}{\mathcal{M}^2-4^C}. \tag{6}$$

According to Fig. 5, the approximated CCC curves are very close to those calculated from explicit equations [16] for 4QAM, 16QAM and 64QAM. Note that conventionally, (4) only defines the capacity for transmitting M_2 at the near receiver. However, in our proposed DN-SR system, both M_1 and M_2 are required by Receiver 2 to re-construct the original data. Therefore, one more rate constraint equation is needed with respect to the transmission of M_1 over Link 2, which follows (3), where γ_1 should be replaced by γ_2. Details can be found in the next section.

2) Successive Refinement: The principle of successive refinement is introduced in this sub-section, considering a case of generating two descriptions from one source. According to the diagram in Fig. 2, $\hat{X}_1$ is recovered by Receiver 1, depending on the basic description M_1 with $d(X,\hat{X}_1) = D_1$, and $\hat{X}_0$ is recovered by Receiver 2, depending on both M_1 and M_2 with $d(X,\hat{X}_0) = D_0$, where M_2 is a refinement of M_1, and $d(\cdot)$ denotes the Hamming distortion measure. Obviously, $D_0 \leq D_1$ due to the refinement effect by M_2. However, there is no stand-alone decoder for M_2. The rates R_{s_1} and R_{s_2} for transmitting M_1 and M_2, respectively, should satisfy the following EI Gamal-Cover inner bound by [9]

$$R_{s_1} \geq I(X;\hat{X}_1), \tag{7}$$

$$R_{s_1}+R_{s_2} \geq I(X;\hat{X}_0,\hat{X}_1). \tag{8}$$

Let $h(x) = -x\log_2(x)-(1-x)\log_2(1-x)$ denote the binary entropy function, (7) and (8) become

$$\begin{aligned} R_{s_1} &\geq H(X) - H(X|\hat{X}_1) \\ &= 1-h(D_1) \\ &= R_d(D_1), \end{aligned} \tag{9}$$

and

$$\begin{aligned} R_{s_1}+R_{s_2} &\geq I(X;\hat{X}_0) + I(X;\hat{X}_1|\hat{X}_0) \\ &= H(X) - H(X|\hat{X}_0) \\ &= 1-h(D_0) \\ &= R_d(D_0), \end{aligned} \tag{10}$$

respectively, where $I(X;\hat{X}_1|\hat{X}_0) = 0$ because $X \to \hat{X}_0 \to \hat{X}_1$ forms a Markov chain.

III. Outage Probability Analysis

The outage probability of a lossy transmission system is defined by the probability that the re-constructed source at the receiver does not meet the required distortion [30]. Specifically, for the proposed DN-SR, outage happens if either Receiver 1 or Receiver 2 can not satisfy the maximum acceptable distortions D_1 or D_0, pre-defined by the system. The maximum allowed rate for error-free transmission can be converted to the instantaneous SNR threshold, as in [18]. According to (3)-(5), given the instantaneous SNRs γ_1 and γ_2, the maximum achievable rates for transmitting M_i over Link j, $j=\{1,2\}$, denoted by R_{s_i,l_j}, can be expressed by

$$\text{Link1}: R_{s_1,l_1} \leq \frac{1}{R_{c_1}}\log_2\left[\frac{1+\frac{\alpha\gamma_1}{\tilde{\alpha}\gamma_1+1}}{\sqrt{1+\frac{\alpha^2\gamma_1^2}{(1+\tilde{\alpha}\gamma_1)^2\mathcal{M}^2}}}\right], \tag{11}$$

$$\text{Link2}: R_{s_1,l_2} \leq \frac{1}{R_{c_1}}\log_2\left[\frac{1+\frac{\alpha\gamma_2}{\tilde{\alpha}\gamma_2+1}}{\sqrt{1+\frac{\alpha^2\gamma_2^2}{(1+\tilde{\alpha}\gamma_2)^2\mathcal{M}^2}}}\right], \tag{12}$$

$$R_{s_2,l_2} \leq \frac{1}{R_{c_2}}\log_2\left[\frac{1+\tilde{\alpha}\gamma_2}{\sqrt{1+\frac{\tilde{\alpha}^2\gamma_2^2}{\mathcal{M}^2}}}\right], \tag{13}$$

where the joint probability density function (PDF) of γ_1 and γ_2 can be expressed by $p(\gamma_1,\gamma_2)=p(\gamma_1)p(\gamma_2)$, and

$$p\left(\gamma_i\right)=\frac{1}{\Gamma_i}\exp(-\frac{\gamma_i}{\Gamma_i}),\quad i=1,2 \tag{14}$$

with Γ_j denoting the average received SNR of Link j. It is found that three rate constraints are defined in in (11)-(13), which are used for different receiving strategies. Specifically, only R_{s_1,l_1} for transmitting M_1 is considered in Link 1, as shown in (11), while R_{s_1,l_2} and R_{s_2,l_2} are considered for transmitting M_1 and M_2 in Link 2, as shown in (12) and (13). The transmission is successful if the maximally allowed R_{s_1,l_1}, R_{s_1,l_2} and R_{s_2,l_2} all satisfy the distortion requirements defined by different schemes. Note that the achievability of R_{s_1,l_1} is independent of R_{s_1,l_2} and R_{s_2,l_2}, but R_{s_1,l_2} and R_{s_2,l_2} are further constrained by $R_{c_1}R_{s_1,l_2}+R_{c_2}R_{s_2,l_2}\leq C(\gamma_2)$, because they are both determined by γ_2. Given a specific power allocation ratio α, the outage probabilities of Link 1 and Link 2 transmissions can be independently calculated, based on the assumption of independent block Rayleigh fading channels.

A. Proposed DN-SR System

The outage probability of the proposed DN-SR system is derived in this sub-section. For Link 1, successful transmission only depends on the achievability of R_{s_1,l_1}, as required in (9). For Link 2, since the correlation between M_1 and M_2 is eliminated, the constraint of the sum rate $(R_{s_1,l_2}+R_{s_2,l_2})$, as required in (10), can be individually satisfied. In other words, even though the transmission of X is lossy, transmissions of its lossy descriptions M_1 and M_2 must be loss-less. Therefore, successful transmission of the entire system is guaranteed if the maximum achievable values of R_{s_1,l_1}, R_{s_1,l_2} and R_{s_2,l_2} are all within the admissible rate regions. Let $P_{a_1,sr}$ and $P_{a_2,sr}$ denote the probabilities that the required rates of descriptions are achieved at Receiver 1 and Receiver 2, respectively, we have

$$P_{a_1,sr}=\text{Prob}\left[R_{s_1,l_1}\geq R_d(D_1)\right], \tag{15}$$

$$P_{a_2,sr}=\text{Prob}\left[R_{s_1,l_2}\geq R_d(D_1),\, R_{s_2,l_2}\geq \Delta R_d(D_0,D_1)\right], \tag{16}$$

where the rate distortion function $R_d(\cdot)$ assumes Hamming distortion measure [15] [31] with

$$R_d(D)=\begin{cases}h(e)-h(D), & 0\leq D\leq \min\{e,1-e\}\\ 0, & D>\min\{e,1-e\}\end{cases} \tag{17}$$

where $e=0.5$ for BSS. Note that $\Delta R_d(D_0,D_1)=R_d(D_0)-R_d(D_1)$, which indicates the connection between two transmit rates of the descriptions via successive refinement coding. Then, the outage probability of DN-SR can be expressed as $P_{out,sr}=1-P_{a_1,sr}P_{a_2,sr}$. By combining (11)-(13) and (15)-(16), the rate constraints can be converted to the thresholds of instantaneous received SNRs, given by

$$P_{a_1,sr}=\text{Prob}\left[\gamma_1\geq\frac{z_1}{\alpha+\alpha z_1-z_1}\right], \tag{18}$$

$$P_{a_2,sr}=\text{Prob}\left[\gamma_2\geq\max\left(\frac{z_1}{\alpha+\alpha z_1-z_1},\,\frac{z_2}{1-\alpha}\right)\right], \tag{19}$$

where $z_1=Z(R_d(D_1)R_{c_1})$ and $z_2=Z(\Delta R_d(D_0,D_1)R_{c_2})$, according to (6). The value range of α which guarantees γ_1 and γ_2 being positive is given by

$$\left\{\frac{z_1}{\alpha+\alpha z_1-z_1}>0,\ \frac{z_2}{1-\alpha}>0\right\}\Rightarrow\alpha>\frac{z_1}{1+z_1}. \tag{20}$$

According to (18), $P_{a_1,sr}$ can be further calculated by integrating $p(\gamma_1)$ over the admissible region as

$$\begin{aligned}P_{a_1,sr}&=\int_{\gamma_1=\frac{z_1}{\alpha+\alpha z_1-z_1}}^{\infty}\frac{1}{\Gamma_1}\exp\left(-\frac{\gamma_1}{\Gamma_1}\right)d\gamma_1\\&=-\exp\left(-\frac{\gamma_1}{\Gamma_1}\right)\Big|_{\frac{z_1}{\alpha+\alpha z_1-z_1}}^{\infty}\\&=\exp\left[\frac{-z_1}{\Gamma_1\left(\alpha+\alpha z_1-z_1\right)}\right].\end{aligned} \tag{21}$$

According to (19), the expression of $P_{a_2,sr}$ is found to depend on either $z_1/(\alpha+\alpha z_1-z_1)$ or $z_2/(1-\alpha)$, which is determined by the parameter α, as

$$\frac{z_1}{\alpha+\alpha z_1-z_1}<\frac{z_2}{1-\alpha}\Rightarrow\alpha>\frac{z_1+z_1z_2}{z_1+z_2+z_1z_2}. \tag{22}$$

Therefore, if $\frac{z_1+z_1z_2}{z_1+z_2+z_1z_2}<\alpha<1$,

$$\begin{aligned}P_{a_2,sr}&=\int_{\gamma_2=\frac{z_2}{(1-\alpha)}}^{\infty}\frac{1}{\Gamma_2}\exp\left(-\frac{\gamma_2}{\Gamma_2}\right)d\gamma_2\\&=-\exp\left(-\frac{\gamma_2}{\Gamma_2}\right)\Big|_{\frac{z_2}{(1-\alpha)}}^{\infty}\\&=\exp\left[\frac{-z_2}{\Gamma_2\left(1-\alpha\right)}\right],\end{aligned} \tag{23}$$

and otherwise

$$\begin{aligned}P_{a_2,sr}&=\int_{\gamma_2=\frac{z_1}{\alpha+\alpha z_1-z_1}}^{\infty}\frac{1}{\Gamma_2}\exp\left(-\frac{\gamma_2}{\Gamma_2}\right)d\gamma_2\\&=-\exp\left(-\frac{\gamma_2}{\Gamma_2}\right)\Big|_{\frac{z_1}{\alpha+\alpha z_1-z_1}}^{\infty}\\&=\exp\left[\frac{-z_1}{\Gamma_2\left(\alpha+\alpha z_1-z_1\right)}\right].\end{aligned} \tag{24}$$

Moreover, the two boundaries of α in (20) and (22) have the following relationship,

$$\frac{z_1}{1+z_1}<\frac{z_1+z_1z_2}{z_1+z_2+z_1z_2}, \tag{25}$$

which is proved in Appendix A. Finally, the closed-form expression of the system outage probability is given by (26). Notice that (26) is a segmented equation, depending on the power allocation ratio α. To be specific, (26a) indicates a range of α such that D_1 can never be achieved at both Receiver 1 and Receiver 2, by detecting M_1. In this case, the system is always in outage, regardless of the detection of M_2 at Receiver 2. (26b) indicates the case where the required γ_2 for achieving the constraint of R_{s_1,l_2} is higher than for achieving the constraint of R_{s_2,l_2} at Receiver 2. Thereby, the term on the left hand side of the max function in (19) is used. (26c) indicates the opposite case of (26b).

$$P_{out,sr} = \begin{cases} 1, & 0 < \alpha \leq \frac{z_1}{1+z_1} \quad (26a) \\ 1 - \exp\left[\dfrac{-z_1\left(\Gamma_1+\Gamma_2\right)}{\Gamma_1\Gamma_2\left(\alpha+\alpha z_1 - z_1\right)}\right], & \frac{z_1}{1+z_1} < \alpha < \frac{z_1+z_1z_2}{z_1+z_2+z_1z_2} \quad (26b) \\ 1 - \exp\left[\dfrac{-z_1}{\Gamma_1\left(\alpha+\alpha z_1 - z_1\right)} + \dfrac{-z_2}{\Gamma_2\left(1-\alpha\right)}\right], & \frac{z_1+z_1z_2}{z_1+z_2+z_1z_2} \leq \alpha < 1 \quad (26c) \end{cases}$$

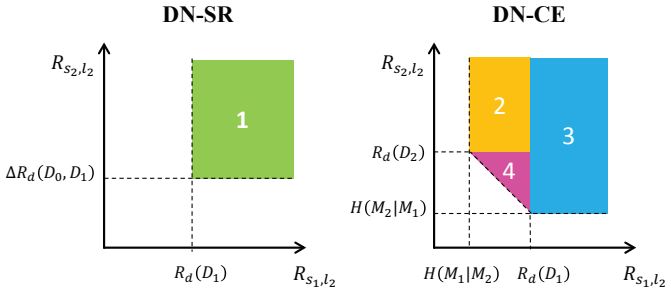


Fig. 6. Admissible rate regions of Link 2 transmission for DN-SR and DN-CE.

B. Conventional DL-NOMA

For the conventional DL-NOMA, $\hat{M}_1$ is not utilized at Receiver 2 for source re-construction. Therefore, only the achievability of R_{s_2,l_2} is considered for Link 2, where $D_2 = D_0$ is defined. Note that without successive refinement coding, $R_d(D_2) \neq \Delta R_d(D_0, D_1)$, because the rates of two descriptions can be independently set. The probability that the conventional DL-NOMA can achieve the admissible rates is given by

$$P_{a,dn} \geq \text{Prob}\left[R_{s_1,l_1} \geq R_d(D_1),\ R_{s_2,l_2} \geq R_d(D_2)\right], \quad (27)$$

where the inequality indicates the fact that Shannon's source-channel separation theorem may not hold in this scenario. Based on (11) and (13), the rate constraints for transmitting descriptions can be converted to that of the instantaneous received SNRs, yielding that

$$P_{a,dn} \geq \text{Prob}\left[\gamma_1 \geq \frac{z_1}{\alpha+\alpha z_1 - z_1},\ \gamma_2 \geq \frac{z_2}{1-\alpha}\right]. \quad (28)$$

The constraint on α which guarantees γ_1 and γ_2 being positive is found the same as in (20). Therefore, if $\frac{z_1}{1+z_1} < \alpha < 1$, the outage probability upper bound of the conventional DL-NOMA is defined by $P_{out,dn} = 1 - P_{a,dn}$, as

$$\begin{aligned} P_{out,dn} =& 1 - \int_{\gamma_1=\frac{z_1}{\alpha+\alpha z_2 - z_1}}^{\infty} \int_{\gamma_2=\frac{z_2}{1-\alpha}}^{\infty} p\left(\gamma_1\right) p\left(\gamma_2\right) d\gamma_1 d\gamma_2 \\ =& 1 - \exp\left(-\frac{\gamma_1}{\Gamma_1}\right)\Big|_{\frac{z_1}{\alpha+\alpha z_2 - z_1}}^{\infty} - \exp\left(-\frac{\gamma_2}{\Gamma_2}\right)\Big|_{\frac{z_2}{1-\alpha}}^{\infty} \\ =& 1 - \exp\left[\frac{-z_1}{\Gamma_1\left(\alpha+\alpha z_1 - z_1\right)} + \frac{-z_2}{\Gamma_2\left(1-\alpha\right)}\right]. \end{aligned} \quad (29)$$

Otherwise, if $0 < \alpha \leq \frac{z_1}{1+z_1}$, $P_{out,dn} = 1$. Note that $P_{out,dn}$ itself is defined as the upper bound, and hence inequality is not needed. The same principle is applied to $P_{out,ce}$ is the following discussion.

C. DL-NOMA with Correlation Exploitation

Since the transmission of Link 1 for DN-CE is kept the same with DN-SR and the conventional DL-NOMA, the probability of achieving the admissible rate R_{s_1,l_1} in Link 1 is the same as (15). For the transmission of Link 2, due to the assumption that $D_0 = 0$, the admissible rate region in terms of R_{s_1,l_2} and R_{s_2,l_2} is analyzed by the Slepian-Wolf theorem, which are given by

$$R_{s_1,l_2} \geq H(M_1|M_2), \quad (30)$$
$$R_{s_2,l_2} \geq H(M_2|M_1), \quad (31)$$
$$R_{s_1,l_2} + R_{s_2,l_2} \geq H(M_1, M_2). \quad (32)$$

According to (30)-(32), the required rates for transmitting M_1 and M_2 are allowed to be less than $R_d(D_1)$ and $R_d(D_2)$ in Link 2 transmission, by exploiting the correlation knowledge at Receiver 2. It indicates that DN-CE can achieve a larger admissible rate region of the descriptions than DN-SR, as shown in Fig. 6. It should be noted that the rate region division for DN-CE is not unique, but to merely simplify the mathematical derivation. With the assumptions described above, the admissible rate regions can be expressed by

$$\begin{aligned} P_{a_2,ce} \geq \text{Prob}[&H(M_1|M_2) \leq R_{s_1,l_2} \leq R_d(D_1), \\ &R_{s_2,l_2} \geq R_d(D_2), \end{aligned} \quad (33)$$
$$P_{a_3,ce} \geq \text{Prob}[R_{s_1,l_2} \geq R_d(D_1),\ R_{s_2,l_2} \geq H(M_2|M_1)], \quad (34)$$
$$\begin{aligned} P_{a_4,ce} \geq \text{Prob}[&H(M_1|M_2) \leq R_{s_1,l_2} \leq R_d(D_1), \\ &R_d(D_0) - R_{s_1,l_2} \leq R_{s_2,l_2} \leq R_d(D_2), \end{aligned} \quad (35)$$

where the inequality indicates the non-optimality of Shannon's source-channel separation theorem in this scenario. Similarly, the rate constraints in (33)-(35) can be translated into the instantaneous SNR constraints as follows.

$$P_{a_2,ce} \geq \text{Prob}\left[\max\left(\frac{z_2}{1-\alpha}, \frac{z_1^*}{\alpha+\alpha z_1^* - z_1^*}\right) \leq \gamma_2 \leq \frac{z_1}{\alpha+\alpha z_1 - z_1}\right] \quad (36)$$

$$P_{a_3,ce} \geq \text{Prob}\left[\gamma_2 \geq \max\left(\frac{z_1}{\alpha+\alpha z_1 - z_1}, \frac{z_2^*}{1-\alpha}\right)\right], \quad (37)$$

$$\begin{aligned} P_{a_4,ce} \geq \text{Prob}\Bigg[&\frac{z_1^*}{\alpha+\alpha z_1^* - z_1^*} \leq \gamma_2 \leq \frac{z_1}{\alpha+\alpha z_1 - z_1}, \\ &\frac{z_3}{1-\alpha} \leq \gamma_2 \leq \frac{z_2}{1-\alpha}\Bigg], \end{aligned} \quad (38)$$

where $z_1^* = Z(H(M_1|M_2)R_{c_1})$, $z_2^* = Z(H(M_2|M_1)R_{c_2})$, $z_3 = Z(H(M_1, M_2)R_{c_0} - C(\alpha\gamma_2/(\tilde{\alpha}\gamma_2 + 1)))$. $P_{a_2,ce}$ and $P_{a_3,ce}$ are then derived in (38) and (39), respectively. Since obtaining the closed-form expression of $P_{a_4,ce}$ is too difficult,

it will be calculated by Monte Carlo simulations in the numerical studies. Finally, the outage probability upper bound of the DN-CE system is defined by $P_{out,ce} = 1 - P_{a_1,ce}(P_{a_2,ce} + P_{a_3,ce} + P_{out_4,ce})$. Note that $P_{a_1,ce} = P_{a_1,sr}$ since the Link 1 transmissions are the same for both DN-SR and DN-CE.

IV. Performance Evaluations

A. Accuracy Evaluation

In this sub-section, the accuracy of the closed-from derivation in (26) for DN-SR is evaluated by comparing its numerical calculations to the Monte Carlo simulation results. As shown in Fig. 7, the outage probability curves of three different cases are plotted versus R_{s_1}, where $D_1 = R_d^{-1}(R_{s_1})$, $D_0 = 0$, and therefore $R_{s_2} = 1 - R_{s_1}$. In order to keep the symbol-level frame length the same for S_1 and S_2, $\eta = R_s/R_c$ is also set as a parameter. Moreover, $\Gamma_1 = \Gamma_2 + 3$ (dB) is assumed to separate the near and far receivers. The parameters of the three cases are listed in Table. I.

TABLE I: System paramters for accuracy evaluation

	α	η	$\mathcal{M}$	Γ_1
Case 1	0.4	1.5	4	15
Case 2	0.6	1.0	16	20
Case 3	0.8	0.7	64	25

For each case, the Monte Carlo simulation is performed based on (18) and (19), with 1 million random channel realizations. It can be found that, despite the parameter adjustment shown in Table. I, the outage curves obtained by the closed-form derivation are very close to that obtained by Monte Carlo simulations. Hence, the accuracy of our mathematical derivation of (26) has then been proven. The slightly visible gaps are due to the approximation of (5) when higher order modulation is utilized, which can be also seen in Fig. 5. However, they are too small to impact the comparative studies in this paper.

B. Outage versus R_{s_1} and α

In this sub-section, outage probabilities of the three downlink schemes are compared with different parameter settings. Given a fixed amount of transmission resources, the symbol-level frame lengths are set equal with $\eta = 1$ in the simulations. The same distortion requirements are used for all schemes for the sake of fairness. Specifically, for DN-SR, $D_1 = R_d^{-1}(R_{s_1})$ and $D_0 = 0$, where R_{s_1} is used as the parameter, and obviously $R_{s_2} = 1 - R_{s_1}$. For the conventional DL-NOMA, the same R_{s_1} is used for transmitting M_1. However, $R_{s_2} = 1$ and $D_2 = 0$, which denotes that $M_2 = X$. DN-CE uses the same R_{s_1} and R_{s_2} as the conventional DL-NOMA. Obviously, $H(M_1|M_2) = H(M_1|X) = 0$ and $H(M_2|M_1) = 1 - R_{s_1}$. DN-CE also requires $D_0 = 0$ as DN-SR, since joint source decoding is applied. In addition, $\Gamma_2 = \Gamma_1 + 3$ (dB) is assumed in the simulations to separate the far and near receivers, with $\mathcal{M} = 4$.

The superiority of DN-CE over the conventional DL-NOMA is first verified. As shown in Fig. 8, the outage probabilities of the two schemes are calculated, with R_{s1} and α being the parameters. Moreover, $\Gamma_1 = 15$ dB and $\Gamma_2 = \Gamma_1 + 3$ dB are used in the simulations. According to Fig. 8, DN-CE can always achieve a lower outage probability than the conventional DL-NOMA, except the corner part with large α and small R_{s_1}, resulting the failure of both schemes. Obviously, the performance gain of DN-CE comes from the extended admissible rate region of M_1 and M_2, by exploiting their correlation at the receiver side. The corner part indicates that the allocated transmit power for M_2 is too small to guarantee its successful detection at Receiver 2, regardless of the detection of M_1.

According to Fig. 6, DN-CE achieves a larger admissible rate region than DN-SR. For example, by fixing $D_1 = R_d^{-1}(R_{s_1})$ for the two schemes with $R_{s_1} = 0.65$, DN-SR requires at least a rate pair $(0.65, 0.35)$ for achieving $D_0 = 0$ at Receiver 2. In contrast, DN-CE has more choices of the allowed rate pair, such as the point $(0.55, 0.45)$ within the sub-region indicated by the number 4. However, such observation does not indicate that DN-CE always outperforms DN-SR. As shown in Fig. 9(a), the outage probability surfaces of the two schemes are plotted in three dimension (3D) versus R_{s_1} and α, which intersect according to the parameters. Clearly, with $R_{s_1} = 0.65$, DN-CE could not achieve a lower outage probability than DN-SR for a certain α value range. Hence, despite the advantage of having a larger admissible rate region of the descriptions, DN-CE may not outperform DN-SR in terms of outage probabilities under particular transmission parameters. Mathematical explanations of this issues is detailed in Appendix B.

It should be noted from Fig. 9(a) that DN-CE outperforms DN-SR in low R_{s_1} and high α value ranges, except the flat part where the both schemes fail to satisfy the required distortions. However, the performance gains achieved by DN-CE are not as significant as those achieved by DN-SR in the opposite side of the parameter value ranges. The results in Fig. 9(a) are further projected to a two dimension (2D) plane, as shown in Fig. 9(b). Specifically, the outage probabilities are only versus R_{s_1}, by taking into account all tested values of α. Clearly, for any fixed R_{s_1}, DN-SR is found to always achieve a lower outage probability than DN-CE, by choosing an appropriate value of α. This observation proves the superiority of DN-SR in the practical system design.

C. Performances versus SNR

In this sub-section, explicit outage performances are evaluated versus the average SNR. First of all, with $\alpha = 0.8$, the outage probabilities of the three schemes are plotted versus Γ_1, as shown in Fig. 10. It is found that the proposed DN-SR system achieves the lowest outage probabilities, with roughly 4 dB and 6 dB gains over the DN-CE and the conventional DL-NOMA systems, respectively. On the other hand, by exploiting the correlation knowledge, DN-CE exhibits a better performance than conventional DL-NOMA. However, despite the rate region expansion achieved by DN-CE, the necessity for increasing the channel coding rate leads to a higher outage performance than DN-SR, due to the symbol length constraint. It is also reasonable to observe that the outage probabilities

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.3022400, IEEE Transactions on Communications

$$P_{a_2,ce} \geq \begin{cases} 0 & , \alpha < \frac{z_1}{1+z_1} \quad (39a) \\ \exp\left[\frac{-z_2}{\Gamma_2(\alpha+\alpha z_2 - z_2)}\right] - \exp\left[\frac{-z_1}{\Gamma_2(\alpha+\alpha z_1 - z_1)}\right] & , \frac{z_1}{1+z_1} \leq \alpha < \frac{z_2+z_1^* z_2}{z_1^*+z_2+z_1^* z_2} \quad (39b) \\ \exp\left[\frac{-z_1^*}{\Gamma_2(1-\alpha)}\right] - \exp\left[\frac{-z_1}{\Gamma_2(\alpha+\alpha z_1 - z_1)}\right] & , \frac{z_1^*+z_1^* z_2}{z_1^*+z_2+z_1^* z_2} \leq \alpha < \frac{z_1+z_1 z_2}{z_1+z_2+z_1 z_2} \quad (39c) \\ 0 & ,\alpha \geq \frac{z_1+z_1 z_2}{z_1+z_2+z_1 z_2} \quad (39d) \end{cases}$$

$$P_{a_3,ce} \geq \begin{cases} 0 & , \alpha < \frac{z_1}{1+z_1} \quad (40a) \\ \exp\left[\frac{-z_1}{\Gamma_2(\alpha+\alpha z_1 - z_1)}\right] & , \frac{z_1}{1+z_1} \leq \alpha < \frac{z_1+z_1 z_2^*}{z_1+z_2^*+z_1 z_2^*} \quad (40b) \\ \exp\left[\frac{-z_2^*}{\Gamma_2(1-\alpha)}\right] & , \alpha \geq \frac{z_1+z_1 z_2^*}{z_1+z_2^*+z_1 z_2^*} \quad (40c) \end{cases}$$

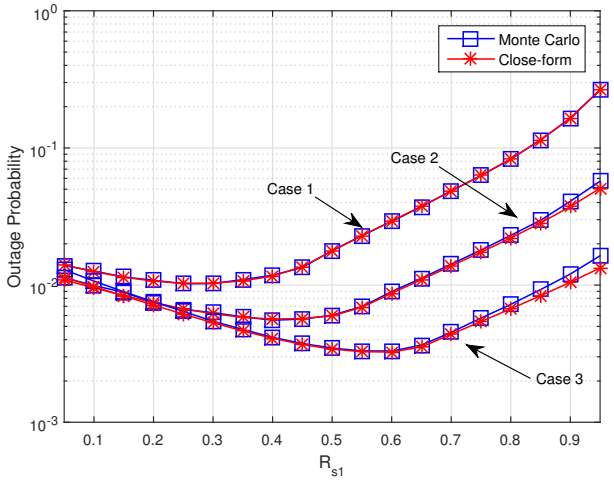


Fig. 7. Accuracy investigation by comparing the closed-form and Monte Carlo simulations of DN-SR.

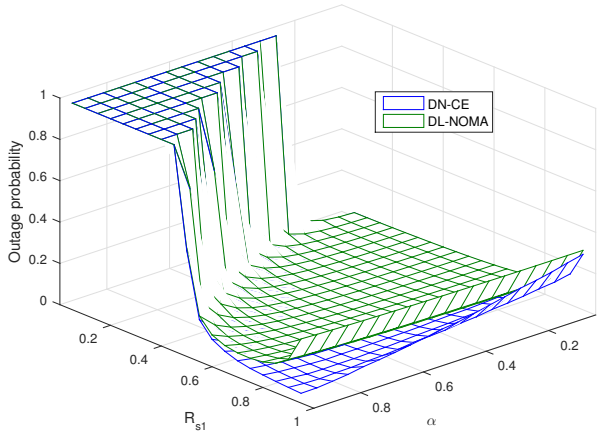


Fig. 8. Outage probability comparisons between DL-NOMA and DN-CE versus R_{s_1} and α.

become larger by increasing R_{s_1} from 0.35 to 0.65, indicating a more strict requirement of D_1.

Additionally, the simulations are reproduced by decreasing α from 0.8 to 0.45, of which results are shown in Fig. 11. It should be noted that with $R_{s_1} = 0.65$, DN-CE achieves a lower outage curve than DN-SR. The main reason is that, the loss in channel coding gain can be fully compensated by the rate region expansion of the source. However, this tendency is not observed with $R_{s_1} = 0.35$. According to the observations of Fig. 10 and Fig. 11, the gaps of outage curves are not sensitive to the investigated average SNR values.

V. Optimal Power Allocation

Based on the closed-form expression of (26), this section aims to minimize the outage probability of DN-SR with an optimal power allocation ratio α, while fixing the other transmission parameters. The problem can be formulated by

$$\begin{aligned} \text{minimize} \quad & P_{out,sr}(\alpha) \\ \text{subject to} \quad & \alpha - 1 \leq 0, \\ & -\alpha \leq 0. \end{aligned} \quad (41)$$

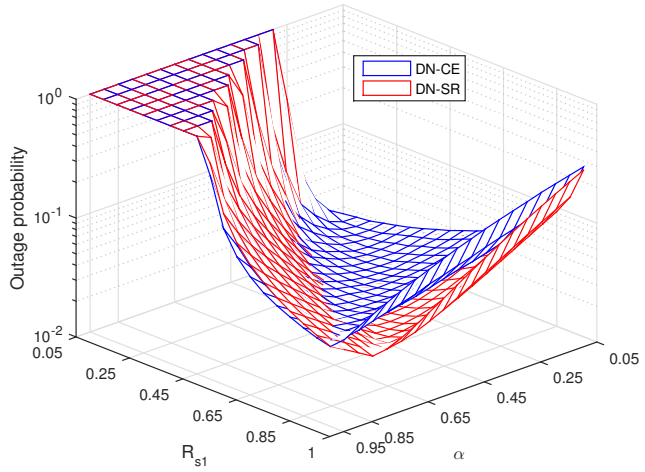
where the function $P_{out,sr}$ contains a single variable of α. Even though $P_{out,sr}$ defined in (26) is segment-wise, the optimal α can still be obtained by solving a convex optimization problem [32]. The convexity proof is detailed in Appendix C.

In Fig. 12, the optimal α values versus R_{s_1} are plotted, with different gaps between Γ_1 and Γ_2 in dB as parameters. It can be seen that the larger the R_{s_1}, the larger the optimal α. Moreover, the increase in the optimal α is found to be slower in a high R_{s_1} value range. Furthermore, if Γ_2 increases, i.e., the Link 2 transmission becomes more reliable, the optimal α is found to increase, and more power can be allocated to the Link 1 transmission. However, the optimal α values with different average SNR gaps becomes very close in high R_{s_1} value range.

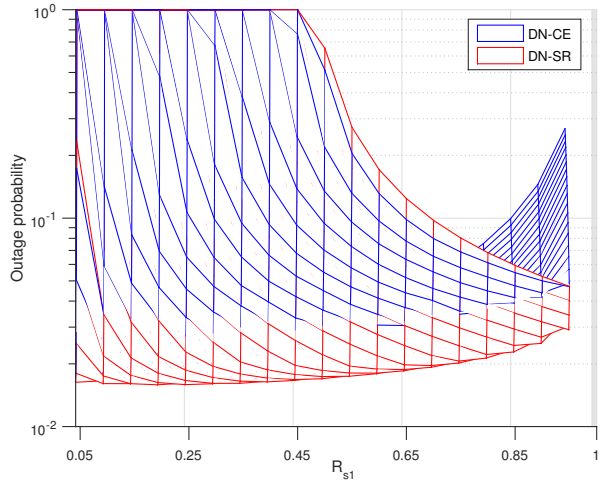
VI. Conclusion

To enable a flexible information delivery to receivers having different QoS requirements, this paper has proposed an efficient lossy transmission of BSS by applying successive refinement to the conventional DL-NOMA, referred to as DN-SR. By jointly considering the broadcast channel capacity and the EI Gamal-Cover inner bound, the admissible rate region of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2020.3022400, IEEE Transactions on Communications



(a) Outage probability in 3D.



(b) Projected outage probability in 2D.

Fig. 9. Outage probability comparisons between DN-SR and DN-CE versus R_{s_1} and α.

DN-SR was identified. As the main contribution of this paper, the system outage probability was derived in closed-form, under block Rayleigh fading assumption. The performance gain achieved by DN-SR has been observed compared to the conventional DL-NOMA. Furthermore, an enhanced DL-NOMA was studied, which exploits the correlation between two descriptions, referred to as DN-CE. Despite a larger admissible rate region with DN-CE for source coding, it could not achieve a lower outage probability than DN-SR, with appropriate parameter adjustment. Finally, the optimal power allocation ratio for DN-SR is determined by solving the convex optimization problem over the derived closed-form expressions. Investigation of more than 2 receivers' scenario is left as a future work.

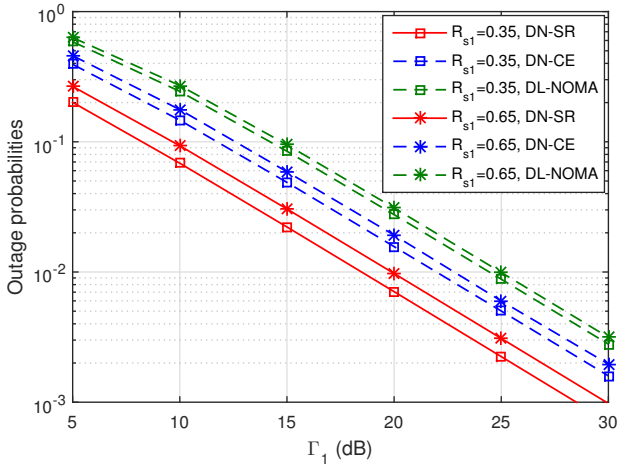


Fig. 10. Outage probability comparisons of three schemes versus Γ_1, with $\alpha = 0.8$.

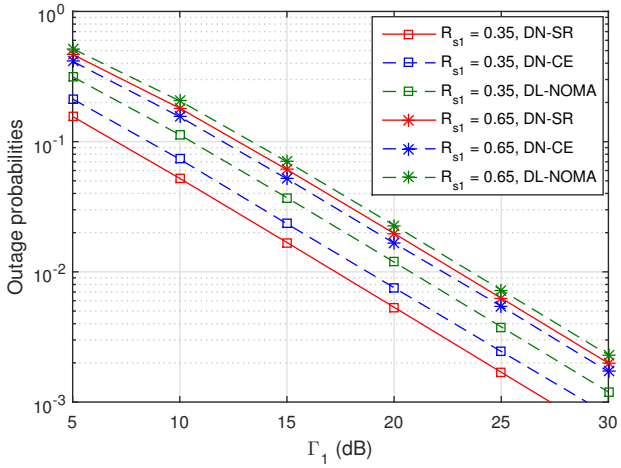


Fig. 11. Outage probability comparisons of three schemes versus Γ_1, with $\alpha = 0.45$.

APPENDIX A
PROOF OF THE INEQUALITY (25)

Since that z_1 and z_2 are both positive, the inequality (22) holds if

$$\begin{aligned} \frac{1}{1+z_1} &< \frac{1+z_2}{z_1+z_2+z_1z_2} \\ \frac{1}{1+z_1} &< \frac{1+z_2}{z_2+z_1(1+z_2)} \\ \frac{1}{1+z_1} &< \frac{1}{\frac{z_2}{1+z_2}+z_1}. \end{aligned} \tag{42}$$

Clearly, (41) holds if $\frac{z_2}{1+z_2} < 1$, and therefore the inequality of (22) has been proven.

APPENDIX B
RELATIONSHIP BETWEEN DN-SR AND DN-CE

According to Fig. 6, the admissible rate region of M_1 and M_2 with DN-SR is covered by that with DN-CE. However, the system performance depends on whether the original source X

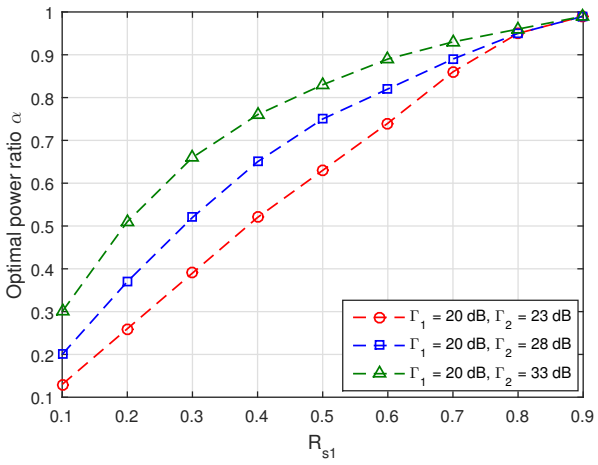


Fig. 12. Optimal power ratio α versus R_{s_1} with different gaps of average SNR.

can be recovered with a required distortion. Given a common distortion D_0, the rate distortion function is given by

$$\begin{aligned} R_d(D_0) &= H(X) - h(D_0) \\ &= H(M_1, M_2) - h(D_0) \\ &= H(M_1) + H(M_2) - I(M_1; M_2) - h(D_0), \end{aligned} \quad (43)$$

where $h(D_0)$ is the binary entropy function. Since DN-SR and DN-CE use different encoding schemes for X, let M_1^{SR} and M_2^{SR} denote the two descriptions for DN-SR, and similarly M_1^{CE} and M_2^{CE} for DN-CE. Due to the fact that M_1^{SR} and M_2^{SR} are independent, while M_1^{CE} and M_2^{CE} are correlated, $I(M_1^{SR}; M_2^{SR}) = 0$ and $I(M_1^{CE}; M_2^{CE}) > 0$. Therefore, according to (42), it is obvious that

$$H(M_1^{SR}) + H(M_2^{SR}) < H(M_1^{CE}) + H(M_2^{CE}), \quad (44)$$

which indicates that DN-CE has to transmit the descriptions with higher rates than DN-SR, in order to achieve the same distortion D_0 for a common source X. In this case, DN-CE must also increase the normalized spectrum efficiency R_c, denoting the product of channel coding rate and modulation order. This is because the total transmission resources are fixed, so the number of modulated symbols within one frame should not change. As a result, given the same SNRs, the description rate pair (R_{s_1}, R_{s_2}) achieved by DN-SR may not be achieved by DN-CE, because of a higher R_c value used in DN-CE. In other words, despite a larger admissible rate region, DN-CE always requires a higher SNR for achieving the same (R_{s_1}, R_{s_2}) with DN-SR.

It is concluded that DN-SR is not a special case of DN-CE, and the superiority of their performances is not dominated by the source rate regions, but depends on the parameters, such as the power allocation ratio and transmitting rates of the descriptions.

Appendix C
Proof of Convexity

The convexity proof of the outage probability expression (26) is provided for DN-SR. For the first segment, $P_{out,sr}(\alpha) \equiv 1$ is defined by (26a), and obviously the optimal α does not take in this segment. For the second segment defined by (26b), the derivative of $P_{out}(\alpha)$ can be given by

$$P'_{out,sr} = \frac{-(\Gamma_1+\Gamma_2)z_1(z_1+1)}{\Gamma_1\Gamma_2\left(\alpha+\alpha z_1 - z_1\right)^2} \exp\left[\frac{-z_1(\Gamma_1+\Gamma_2)}{\Gamma_1\Gamma_2\left(\alpha+\alpha z_1 - z_1\right)}\right], \quad (45)$$

which is always negative. Therefore, $P_{out,sr}(\alpha)$ within the second segment is proved to be a monotonously decreasing function, and the lowest outage is obviously on the right-hand-side boundary of α, and furthermore, the outage curve has to be continuously connected to the third segment. Hence, we only need to find the optimal α in the third segment. The third segment defined by (26c) can be written by a composition of two functions, i.e., $P_{out,sr}(\alpha) = f[g(\alpha)]$, where

$$f(\alpha) = 1 - \exp(\alpha), \quad (46)$$

$$g(\alpha) = \frac{-z_1}{\Gamma_1\left(\alpha+\alpha z_1 - z_1\right)} + \frac{-z_2}{\Gamma_2(1-\alpha)}. \quad (47)$$

The convexity of (26c) requires that the second derivative of $P_{out,sr}(\alpha)$ is non-negative, i.e.,

$$P''_{out,sr}(\alpha) = f''[g(\alpha)]\,g'(\alpha)^2 + f'[g(\alpha)]\,g''(\alpha) \geq 0. \quad (48)$$

It is easy to find that $f(\alpha)$ is a monotonically decreasing concave function, with $f(\alpha)' = f(\alpha)'' = -\exp(\alpha) < 0$, and the condition required in (47) can be written by

$$-g''(\alpha) \geq g'(\alpha)^2, \quad (49)$$

where

$$g'(\alpha) = \frac{z_1(1+z_1)}{\Gamma_1\left(\alpha+\alpha z_1 - z_1\right)^2} - \frac{z_2}{\Gamma_2(1-\alpha)^2}, \quad (50)$$

$$g''(\alpha) = \frac{-2z_1(1+z_1)^2}{\Gamma_1\left(\alpha+\alpha z_1 - z_1\right)^3} + \frac{-2z_2}{\Gamma_2(1-\alpha)^3}. \quad (51)$$

With the parameters assumed in this paper, it is very easy to numerically calculate the equations above, and prove that (48) holds. Therefore $P_{out,sr}(\alpha)$ can be proved to be convex over the third segment, where the optimal α exists.

References

[1] J. Montalban, P. Scopelliti, M. Fadda, E. Iradier, C. Desogus, P. Angueira, M. Murroni, and G. Araniti, "Multimedia multicast services in 5G networks: Subgrouping and non-orthogonal multiple access techniques," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 91–95, Mar. 2018.

[2] C. Regueiro, J. Montalban, J. Barrueco, M. Velez, P. Angueira, Y. Wu, L. Zhang, S. Park, J. Lee, and H. M. Kim, "LDM core services performance in atsc 3.0," *IEEE Trans. Broadcast.*, vol. 62, no. 1, pp. 244–252, Mar. 2016.

[3] L. Zhang, W. Li, Y. Wu, K. Salehian, S. Laflche, Z. Hong, S. Park, H. M. Kim, J. Lee, N. Hur, X. Wang, P. Angueira, and J. Montalban, "Using layered-division-multiplexing to deliver multi-layer mobile services in ATSC 3.0," *IEEE Trans. Broadcast.*, vol. 65, no. 1, pp. 40–52, Mar. 2019.

[4] Z. Ding, M. Peng, and H. V. Poor, "Cooperative non-orthogonal multiple access in 5G systems," *IEEE Commun. Lett.*, vol. 19, no. 8, pp. 1462–1465, Aug 2015.

[5] M. Cheng, Y. Wu, and Y. Chen, "Capacity analysis for non-orthogonal overloading transmissions under constellation constraints," in *Proc. Int. Conf. on Wireless Commun. Signal Process. (WCSP)*, Oct. 2015, pp. 1–5.

[6] L. Zhang, Y. Wu, W. Li, K. Salehian, S. Lafleche, X. Wang, S. I. Park, H. M. Kim, J. Lee, N. Hur, P. Angueira, and J. Montalban, "Layered-division multiplexing: An enabling technology for multicast/broadcast service delivery in 5G," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 82–90, Mar. 2018.

[7] T. Manglayev, R. C. Kizilirmak, Y. H. Kho, N. Bazhayev, and I. Lebedev, "Noma with imperfect sic implementation," in *Proc. IEEE EUROCON 2017, 17th Int. Conf. Smart Technol.*, Jul. 2017, pp. 22–25.

[8] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6010–6023, Aug. 2016.

[9] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. New York, NY, USA: *Cambridge University Press*, 2012.

[10] S. Sesia, G. Caire, and G. Vivier, "Lossy transmission over slow-fading AWGN channels: a comparison of progressive and superposition and hybrid approaches," in *Proc. Int. Symp. Inf. Theory (ISIT)*, Sep. 2005, pp. 224–228.

[11] C. Tian, A. Steiner, S. Shamai, and S. N. Diggavi, "Successive refinement via broadcast: Optimizing expected distortion of a gaussian source over a gaussian fading channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 2903–2918, Jul. 2008.

[12] C. T. K. Ng, D. Gunduz, A. J. Goldsmith, and E. Erkip, "Distortion minimization in gaussian layered broadcast coding with successive refinement," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5074–5086, Nov. 2009.

[13] S. A. Yildirim and M. Yuksel, "Multiple description coding based compress-and-forward for the broadcast relay channel," in *Proc. Int. Symp. on Inf. Theory (ISIT)*, Jul. 2012, pp. 194–198.

[14] M. Benammar, P. Piantanida, and S. Shamai, "Multiple description coding for the compound broadcast channel," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Nov 2014, pp. 566–570.

[15] T. M. Cover and J. A. Thomas, *Elements of information theory*. *John Wiley & Sons*, 2012.

[16] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.

[17] B. Sklar, "Rayleigh fading channels in mobile digital communication systems .I. characterization," *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 90–100, Jul. 1997.

[18] M. Cheng, K. Anwar, and T. Matsumoto, "Outage probability of a relay strategy allowing intra-link errors utilizing slepian-wolf theorem," *EURASIP J. Adv. Signal Process.*, vol. 2013, no. 1, p. 34, Feb. 2013.

[19] X. Zhou, M. Cheng, X. He, and T. Matsumoto, "Exact and approximated outage probability analyses for decode-and-forward relaying system allowing intra-link errors," *IEEE Trans. Wireless Commun.*, vol. 13, no. 12, pp. 7062–7071, Dec. 2014.

[20] J. He, V. Tervo, X. Zhou, X. He, S. Qian, M. Cheng, M. Juntti, and T. Matsumoto, "A tutorial on lossy forwarding cooperative relaying," *IEEE Commun. Surv. Tutorials*, vol. 21, no. 1, pp. 66–87, Aug. 2019.

[21] S. Mohajer, C. Tian, and S. N. Diggavi, "On source transmission over deterministic relay networks," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Jan. 2010, pp. 1–5.

[22] C. Tian, J. Chent, S. N. Diggavi, and S. Shamai, "On source-channel separation in networks," in *Proc. Int. Conf. on Signal Process. & Commun. (SPCOM)*, Jul. 2010, pp. 1–5.

[23] C. Tian, J. Chen, S. N. Diggavi, and S. Shamai, "Optimality and approximate optimality of source-channel separation in networks," in *Proc. Int. Symp. on Inf. Theory (ISIT)*, Jun. 2010, pp. 495–499.

[24] X. Zhou, X. He, M. Juntti, and T. Matsumoto, "Outage probability of correlated binary source transmission over fading multiple access channels," in *Proc. IEEE 16th Int. Workshop on Signal Process. Adv. in Wireless Commun. (SPAWC)*, Jun. 2015, pp. 96–100.

[25] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, Jul. 1973.

[26] K. Anwar and T. Matsumoto, "Accumulator-assisted distributed turbo codes for relay systems exploiting source-relay correlation," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1114–1117, Jul. 2012.

[27] T. Matsumoto, S. Ibi, S. Sampei, and R. Thoma, "Adaptive transmission with single-carrier multilevel BICM," *Proc. IEEE*, vol. 95, no. 12, pp. 2354–2367, Dec. 2007.

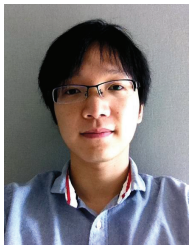
[28] M. Cheng, X. Zhou, K. Anwar, and T. Matsumoto, "Simple relay systems with bicm-id allowing intra-link errors," *IEICE Trans. Commun.*, vol. E95.B, no. 12, pp. 3671–3678, 2012.

[29] W. Li, H. Yang, and D. Yang, "Approximation formulas for the symmetric capacity of m-ary modulations," *Sciencepaper Online*, Apr. 2007.

[30] P. Lu, X. Zhou, and T. Matsumoto, "Outage probabilities of orthogonal multiple-access relaying techniques with imperfect source-relay links," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, Apr. 2015.

[31] L. D. Davisson, "Rate-distortion theory and application," *Proc. IEEE*, vol. 60, no. 7, pp. 800–808, Jul. 1972.

[32] S. Boyd and L. Vandenberghe, *Convex optimization*. *Cambridge University Press*, 2004.



Meng Cheng (S'12-M'16) received the B.Eng. degree in telecommunication engineering from the Anhui University of Technology, Anhui, China, in 2009, the M.Sc. degree (with Distinction) in wireless communications from the University of Southampton, Southampton, U.K., in 2010, and the Ph.D. degree in information science from the Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan, in 2014. He served as a 5G research engineer in the Shanghai research center, Huawei Technologies Co., Ltd from 2014 to 2017. After that, he returned to JAIST as a post-doc researcher. His research interests are network information theory, non-orthogonal multiple access (NOMA), iterative coding/decoding and wireless geolocation techniques.

Wensheng Lin (S'17-M'20) received the B.Eng. degree in communication engineering, and the M.Eng. degree in electronic and communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2013 and 2016, respectively, and the Ph.D. degree in information science from Japan Advanced Institute of Science and Technology (JAIST), Ishikawa, Japan, in 2019. He is currently an Associate Professor with the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. His research interests include network information theory, distributed source coding, and Age of Information.

Tad Matsumoto (S'84-M'98-F'10) received the B.S. and M.S. degrees in electrical engineering under the mentorship of Prof. S.-I. Takahashi, and the Ph.D. degree in electrical engineering under the supervision of Prof. M. Nakagawa from Keio University, Yokohama, Japan, in 1978, 1980, and 1991, respectively.

He joined Nippon Telegraph and Telephone Corporation (NTT) in 1980, where he was involved in a lot of research and development projects, all for mobile wireless communications systems. In 1992, he transferred to NTT DoCoMo, where he researched on code-division multiple-access techniques for mobile communication systems. In 1994, he transferred to NTT America, where he served as a Senior Technical Advisor of a joint project between NTT and NEXTEL Communications. In 1996, he returned to NTT DoCoMo, where he served as the Head of the Radio Signal Processing Laboratory until 2001. He researched on adaptive signal processing, multiple-input multiple-output turbo signal detection, interference cancellation, and space-time coding techniques for broadband mobile communications. In 2002, he moved to the University of Oulu, Finland, where he served as a Professor with the Centre for Wireless Communications. In 2006, he served as a Visiting Professor with the Ilmenau University of Technology, Ilmenau, Germany, funded by the German MERCATOR Visiting Professorship Program. Since 2007, he has been serving as a Professor with the Japan Advanced Institute of Science and Technology, Japan, while also keeping a cross-appointment position with the University of Oulu.

Prof. Matsumoto has led a lot of projects funded by Academy-of-Finland, European FP7, and Japan Society for the Promotion of Science as well as by Japanese private companies. He has been appointed as a Finland Distinguished Professor from 2008 to 2012, funded by the Finnish National Technology Agency (Tekes) and Finnish Academy, under which he preserves the rights to participate in and apply to European and Finnish national projects. He was a recipient of the IEEE VTS Outstanding Service Award in 2001, the Nokia Foundation Visiting Fellow Scholarship Award in 2002, the IEEE Japan Council Award for Distinguished Service to the Society in 2006, the IEEE Vehicular Technology Society James R. Evans Avant Garde Award in 2006, the Thuringen State Research Award for Advanced Applied Science in 2006, the 2007 Best Paper Award of Institute of Electrical, Communication, and Information Engineers of Japan in 2008, the Telecom System Technology Award by the Telecommunications Advancement Foundation in 2009, the IEEE Communication Letters Exemplary Reviewer in 2011, the Nikkei Wireless Japan Award in 2013, the IEEE VTS Recognition for Outstanding Distinguished Lecturer in 2016, and the IEEE TRANSACTIONS ON COMMUNICATIONS Exemplary Reviewer in 2018. He is a Member of IEICE. He has been serving as an IEEE Vehicular Technology Distinguished Speaker since 2016.