Source and Channel Coding for Cooperative Relaying

Deniz Gündüz, *Student Member, IEEE,* and Elza Erkip, *Senior Member, IEEE*

***Abstract*— User cooperation is a powerful tool to combat fading and increase robustness for communication over wireless channels. Although it is doubtless a promising technique for enhancing channel reliability, its performance in terms of average source distortion is not clear since source-channel separation theorem fails under the most common non-ergodic slow fading channel assumption, when channel state information is only available at the receiving terminals. This work sheds some light on the end-to-end performance of joint source-channel coding for cooperative relay systems in the high SNR regime. Considering distortion exponent as a figure of merit, we propose various strategies for cooperative source and channel coding that significantly improve the performance compared to the conventional scheme of source coding followed by cooperative channel coding. We characterize the optimal distortion exponent of a full-duplex relay channel for all bandwidth ratios. For the half-duplex relay channel, we provide an upper bound which is tight for small and large bandwidth ratios. We consider the effect of correlated side information on the distortion exponent as well.**

***Index Terms*— Block fading channel, cooperative transmission, distortion exponent, diversity-multiplexing tradeoff, relay channel.**

I. Introduction

WIRELESS communication technology has entered a new era where it evolved from a system offering mainly voice service to one that provides services with rich multimedia content. The increased demand for different services at the application layer results in higher transmission rate and reliability requirements at the physical layer. However, high data rate and high reliability are two conflicting design parameters. Accordingly, we need an end-to-end performance measure for the overall system.

In this paper, we are interested in transmission of a continuous amplitude (analog) source over a quasi-static fading relay channel within the delay requirement of the underlying application. The performance measure we consider is the end-to-end average distortion between the source and its reconstruction at the destination. As we argued in detail in [1], Shannon's source-channel separation theorem does not hold when the channel state information (CSI) is not available at the transmitters since the channel is no more ergodic. Therefore the system we consider requires a joint source-channel code design for optimal end-to-end performance. We argue here that, when we are interested in the high signal-to-noise ratio (SNR) behavior of the average distortion, in most cases it suffices to consider separate source and channel encoders and decoders as well as cooperative strategies whose parameters are optimized jointly. We specifically consider distortion exponent (Δ) [2] as our performance measure, which is defined as the exponential decay rate of the end-to-end expected distortion in the high SNR limit.

Bandwidth ratio, which is the ratio between the channel and source bandwidths, or the number of channel uses per source sample, plays a crucial role in the achievable system performance. In [1], [12]–[14], we consider a MIMO block fading channel, and analyze its distortion exponent as a function of the bandwidth ratio. We observe that the diversity-multiplexing tradeoff (DMT) is a useful tool in characterizing the distortion exponent, and propose various strategies that utilize the DMT curve of the MIMO channel to improve the achievable distortion exponent. The three schemes proposed in [1] are *layered source coding with progressive transmission* (LS), its hybrid digital-analog extension called *hybrid LS* (HLS), and *broadcast strategy with layered source* (BS). These schemes adapt to channel variations by utilizing layered compression followed by variable rate channel coding that provides unequal error protection. Furthermore, [1] establishes a close relationship between the maximal diversity of the system and its highest achievable distortion exponent. For a MIMO system with M_t transmit and M_r receive antennas and L fading blocks, where the highest possible diversity gain is LM_tM_r, the distortion exponent is also bounded by LM_tM_r and this bound is achievable provided that the bandwidth ratio is high enough.

Our goal in this paper is to characterize the distortion exponent of cooperative relay systems for all bandwidth ratios. User cooperation [3], [4] is a popular spatial diversity technique which utilizes relays to increase diversity. Simple yet effective cooperation schemes such as amplify-and-forward (AF) and decode-and-forward (DF) [5] can result in diversity gains similar to multiple antenna transmission. On the other hand, the full DMT of cooperative transmission might not follow the MIMO tradeoff or the best known achievable strategies fall short of MIMO [6], [7]. Also, for a cooperative system, it is not clear how to utilize the best DMT achieving relaying strategies and combine them with layered source coding in a manner similar to [1] in order to improve the distortion exponent. In the half-duplex relay case, where the relay cannot receive and transmit simultaneously, matching of layered source coding to different channel transmission techniques becomes even more challenging. For example, as we show in Section III,

The material in this paper was presented in part at IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC), New York, NY, June 2005.

This work is partially supported by NSF grant No. 0430885 and No. 0635177.

The authors are with Department of Electrical and Computer Engineering, Polytechnic University, 11201, Brooklyn, NY. Email: dgundu01@utopia.poly.edu, elza@poly.edu

commonly used simple cooperation techniques, such as DF and AF, may achieve a lower distortion exponent than direct transmission (DT) if the bandwidth ratio is not large enough. This shows that the benefit of user cooperation in terms of end-to-end expected distortion is not obvious, and the optimal design requires a careful cross-layer approach.

In this paper, we propose various joint source-channel coding strategies based on layered source coding followed by different channel coding/cooperation schemes and study their distortion exponents. As in MIMO, DMT becomes an effective tool for characterizing the distortion exponent of cooperative systems and maximum diversity serves a bound for the maximum distortion exponent. For half-duplex relaying, we investigate the effect of source layering and illustrate how cooperating on important source layers while sending the rest through direct transmission can increase the distortion exponent. For a single relay system, we find the optimal distortion exponent when the bandwidth ratio is smaller than 1 or greater than 4. When the relay terminal is able to receive and transmit simultaneously, i.e., with a full-duplex relay, we show that BS strategy with block Markov superposition encoding, decode-and-forward relaying and backward decoding at the destination is optimal in terms of distortion exponent for all bandwidth ratios.

Distortion exponent analysis for source transmission over non-ergodic fading channels has gained interest in the research community [1], [2], [8] - [16]. While these papers concentrate on the high SNR behavior, [17] explores the LS scheme for MIMO systems at finite SNR values. For point-to-point communication, we extend the analysis of BS scheme to arbitrary SNR values in [18], [19].

The references [11], [12] contain our initial results on cooperative source transmission over fading channels. In a more practical direction, we also study optimal bit allocation among source coding, channel coding and cooperation for coded cooperative systems in order to realize the end-to-end performance gains through joint source-channel cooperation [20], [21]. In [20], we consider transmission of a Gaussian source over a quasi-static fading relay channel using RCPC coded user cooperation. Later in [21], we utilize these techniques for a video source and show possible gains in decoded video quality for a practical multimedia communication system. Kwasinski et al. [22]- [24] also consider bit allocation among the source and the relay terminals in order to improve end-to-end average distortion for cooperative systems, however, their model is limited as it does not incorporate fading.

The paper is organized as follows: We introduce the system model and definitions in Section II. In Section III, we analyze the distortion exponent of a single rate system and provide an upper bound for the achievable distortion exponent. Then, in Section IV, we consider a half-duplex relay terminal, propose several strategies based on layered compression and find their corresponding achievable distortion exponents. We characterize the optimal distortion exponent of a single half-duplex relay channel for bandwidth ratios $b \leq 1$ and $b \geq 4$. In Section V we find the optimal distortion exponent for the full-duplex relay at all bandwidth ratios. Section VI is devoted to the analysis of distortion exponent when a correlated side information is available at the destination and the relay terminals. Finally, Conclusion and Appendices follow.

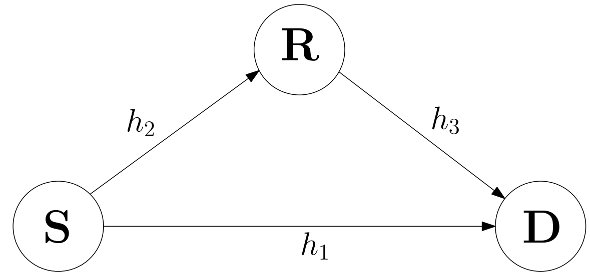

Fig. 1. Illustration of the cooperative system model.

II. SYSTEM MODEL

We assume that a continuous amplitude, memoryless, zero mean, unit variance complex Gaussian source, $\{s_k\}_{k=1}^{\infty}$ is available at the source terminal (S) who wishes to transmit this source to a destination terminal (D) using a relay (R) (see Fig. 1). We consider single-letter squared error distortion, which is defined for length-K source and reconstruction vectors $\mathbf{s} = (s_1, \ldots, s_K)$ and $\hat{\mathbf{s}} = (\hat{s}_1, \ldots, \hat{s}_K)$, respectively, as

$$D(\mathbf{s}, \hat{\mathbf{s}}) = \frac{1}{K} \sum_{k=1}^{K} d(s_k, \hat{s}_k), \quad (1)$$

where $d(s_k, \hat{s}_k) = (s - \hat{s})^2$. We use the corresponding distortion-rate function $D(R) = 2^{-R}$ [25], where R bits per source sample is the source coding rate. It will later be evident that, for our results to hold, the only requirement on the rate-distortion function is the exponential dependence on the rate, which holds for many general source distributions, scalar quantizers [26], many practical encoder approximations [27], and in the case of high resolution [28]. We will also extensively utilize the successive refinability of the underlying Gaussian source [29] in our analysis. However, [30] suggests that, all sources are successively refinable with a limited constant rate loss. Hence as argued in detail in [1], our results can be extended to more general source distributions.

The delay requirements imposed on the system design translate into our model in terms of a limited bandwidth ratio among the channel and the source bandwidths. We assume that K source samples are transmitted in N channel uses, corresponding to a bandwidth ratio of b defined as

$$b = N/K. \quad (2)$$

Note that, N and K are large enough to approach the instantaneous capacity and the rate-distortion bounds.

We consider a single relay assisting the communication between the source and the destination terminals. Each terminal has a single antenna. Links among the terminals are modeled as quasi-static fading with independent Rayleigh distributions. The channel coefficients, denoted by h_1, h_2, and h_3 are independent, circularly symmetric complex Gaussian with variance $1/2$ in each dimension. The quasi-static assumption leads to constant fading for N channel uses. Fading coefficients are known at the corresponding receivers, but not at the transmitters. We also assume an independent additive noise at each receiver modeled as circularly symmetric complex Gaussian with variance $1/2$ in each dimension. Each transmitter has the same power constraint and hence all average received SNRs are equal. Since our analysis will focus on high SNR exponents, considering path loss and asymmetric systems will not change our results. The relay, depending on the physical system constraints, might operate on the half-duplex or the full-duplex modes. A half-duplex relay cannot transmit and receive simultaneously, while a full-duplex relay can.

Our performance measure is the end-to-end expected distortion (ED), which is the expectation of $d(\mathbf{s}, \hat{\mathbf{s}})$ over the source samples, the channel fading and noise distributions. ED depends not only on the source characteristics, the channel model, the distortion metric, and the power constraint of the transmitters, but also on the joint compression, channel coding and transmission techniques used. As discussed in Section I, Shannon's source-channel separation theorem does not hold in our model due to slow fading and lack of CSI at the transmitters. The problem of finding the minimum possible expected distortion for any given channel SNR is an open problem. However, our main focus is on the high SNR behavior of the expected distortion. To capture this behavior, we define the *distortion exponent* [2] as

$$\Delta = -\lim_{\mathrm{SNR}\to\infty} \frac{\log \mathrm{ED}}{\log \mathrm{SNR}}. \tag{3}$$

We want to find the maximum achievable distortion exponent of this cooperative system for a given bandwidth ratio b. Note here the similarity between distortion exponent and diversity gain which is defined as the exponential decay rate of the outage probability $P_{out}(\mathrm{SNR})$, as a function of the average SNR. In general, for a family of codes with rate $R = r\log\mathrm{SNR}$, r is defined as the multiplexing gain of the family, and

$$d(r) = -\lim_{\mathrm{SNR}\to\infty} \frac{\log P_{out}(\mathrm{SNR})}{\log \mathrm{SNR}} \tag{4}$$

as the diversity advantage [31]. The diversity gain $d^*(r)$ is defined as the supremum of the diversity advantage over all possible code families with multiplexing gain r. The relationship between the distortion exponent Δ and $d^*(r)$, also known as the DMT, will be clarified in the next section.

III. Single Rate Source and Channel Coding and Distortion Exponent Upper Bound

In this section, we first illustrate the distortion exponent analysis for single rate direct transmission from the source to the destination, and compare it with simple cooperation protocols. Our results show that, depending on the bandwidth ratio, additional diversity provided by cooperation may not necessarily lead to an improvement in the distortion exponent. Hence a careful design of compression and channel coding strategies, as well as cooperation protocols is essential for improving the distortion exponent. We next provide an upper bound on the distortion exponent, and motivate layered source compression to achieve this upper bound.

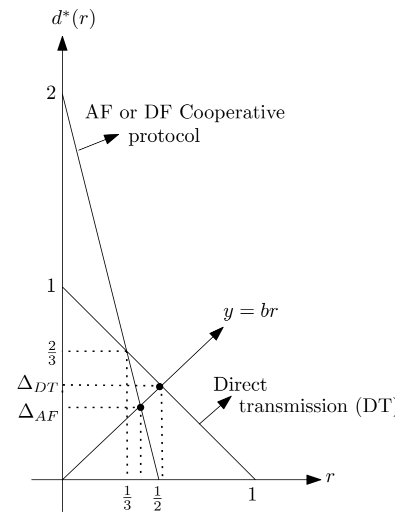

Fig. 2. A geometric interpretation illustrating the optimal multiplexing gain and distortion exponent for a single layer source-channel coding system.

For single rate source and channel coding, we assume a channel code of rate R bits per channel use. The corresponding source coding rate is $NR/K = bR$ bits per source sample. The rate R can be optimized to obtain the single rate distortion exponent.

For bandwidth ratio b, we denote the distortion exponent achieved by single rate source and channel coding with direct transmission (DT) by $\Delta_{DT,1}$, and by simple AF and DF cooperation with a half-duplex relay by $\Delta_{AF,1}$ or $\Delta_{DF,1}$, respectively.

Theorem 3.1: For single rate source and channel coding, we have the following distortion exponents for direct transmission, AF and DF relaying.

$$\Delta_{DT,1} = \frac{b}{b+1}, \tag{5}$$

$$\Delta_{DF,1} = \Delta_{AF,1} = \frac{2b}{b+4}. \tag{6}$$

Proof: We follow an approach similar to [1], [8], where the expected distortion for single rate transmission is written as

$$\mathrm{ED}(R, \mathrm{SNR}) = (1 - P_{out}(R, \mathrm{SNR}))\cdot D(bR) + P_{out}(R, \mathrm{SNR}), \tag{7}$$

where $P_{out}(R, \mathrm{SNR})$ is the outage probability at channel rate R and average received signal to noise ratio SNR, while $D(R)$ is the distortion-rate function of the source. We assume that, in case of an outage, the receiver simply outputs the average value of the source distribution, hence obtains the highest possible distortion of 1. In order to achieve a vanishing expected distortion with increasing SNR, we need to increase rate R with SNR. Scaling R faster than $O(\log\mathrm{SNR})$ would result in outage with probability 1, since the instantaneous channel capacity scales as $\log\mathrm{SNR}$. Thus we assign $R = r\log\mathrm{SNR}$, where $r \in [0, 1]$ is the multiplexing gain. Then

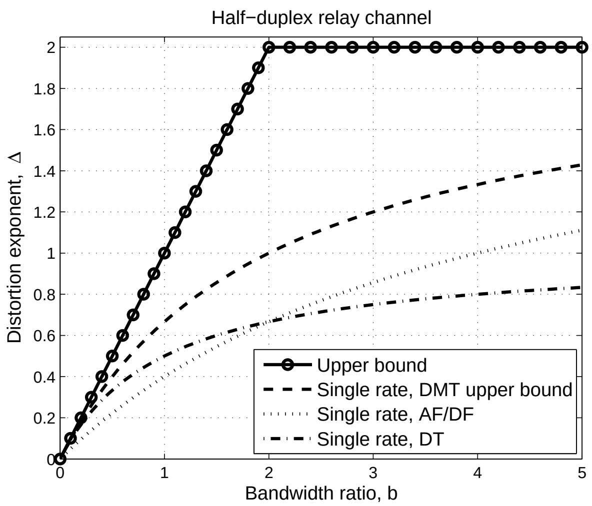

Fig. 3. Distortion exponent of single rate transmission with direct transmission and simple AF/DF cooperation.

the high SNR approximation [1] of Eqn. (7) is given by

$$\begin{aligned} \text{ED} &\doteq D(bR) + P_{out}(R), \\ &\doteq \text{SNR}^{-br} + \text{SNR}^{-d^*(r)}. \end{aligned} \tag{8}$$

where $d^*(r)$ is the diversity gain of the underlying transmission scheme at multiplexing gain r.

Highest distortion exponent is achieved when the two exponents in (8) are equal. Hence the optimal multiplexing gain r^* satisfies

$$br^* = d^*(r^*), \tag{9}$$

which is also equal to the optimal distortion exponent. Using the DMT for DT, $d^*(r) = 1 - r$, we obtain the optimal multiplexing gain and the corresponding optimal distortion exponent as $r^* = 1/(1+b)$, and $\Delta_{DT,1} = b/(1+b)$, respectively. For AF and DF protocols with a half-duplex relay, the DMT is given by $d^*(r) = 2(1-2r)$ [5], which leads to the optimal multiplexing gain and the optimal distortion exponent $r^* = 2/(4+b)$, and $\Delta_{AF,1} = 2b/(4+b)$, respectively. ■

Figure 2 shows a geometric illustration of how the optimal multiplexing gain and the corresponding distortion exponent for Theorem 3.1 can be found. We illustrate the DMT curves of DT, AF and DF protocols together with the $y = br$ line. We observe from (9) and the figure that, the optimal multiplexing gain and the distortion exponent are characterized by the intersection of the $y = br$ line with the DMT curve.

We next provide an upper bound to the distortion exponent.

Theorem 3.2: The distortion exponent of a single relay cooperative system can be upper bounded by

$$\Delta_{UB} = \min\{b, 2\}. \tag{10}$$

Proof: Suppose the source sequence is non-casually available at the relay. In this case, we obtain a 2×1 MISO system, whose distortion exponent upper bound is given in [1]. ■

[1]The symbol $\doteq$ denotes exponential equality, and by $f(\text{SNR}) \doteq \text{SNR}^x$ we mean that $\lim_{\text{SNR}\to\infty} \frac{\log f(\text{SNR})}{\log \text{SNR}} = x$. Exponential inequalities $\dot{\leq}$ and $\dot{\geq}$ are defined similarly.

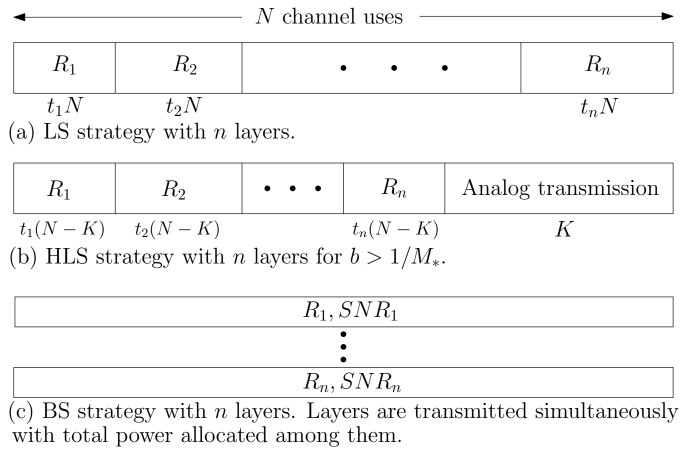

Fig. 4. Illustration of various transmission strategies analyzed in this paper.

Fig. 3 illustrates the distortion exponents of single rate coding for DT and AF/DF protocols along with the upper bound. We note that for $b < 2$, DT results in higher distortion exponent than AF/DF relaying. This can also be deduced from Fig. 2. Not surprisingly, when the system is limited in terms of channel uses per source sample, channel bandwidth becomes a valuable asset and using it for cooperation in order to improve reliability does not pay off. However, for $b > 2$, simple cooperation distortion exponent exceeds that of DT and as $b \to \infty$, AF/DF doubles the distortion exponent compared to DT.

Fig. 3 also shows that the gap between the achievable schemes and the upper bound is significant. In order to see that this gap is not only due to the weakness of the cooperative protocols used, but also arises because of single rate coding, we consider single rate transmission operating on the DMT upper bound for a single relay system, which is equivalent to 2×1 MISO DMT. Using $d^*(r) = 2(1-r)$ yields $\Delta = 2b/(b+2)$ as an upper bound to the single rate distortion exponent of cooperative relay systems. Fig. 3 illustrates that, although this upper bound is above what we have achieved through direct transmission and simple AF or DF for all bandwidth ratios, the gap between the single rate bound and the upper bound of Theorem 3.2 is still significant. This suggests that, along with more sophisticated cooperation protocols, source-channel coding schemes that provide adaptation to the channel variations are essential for improved distortion exponent. Motivated by our results for MIMO systems [1], in the following sections we consider compressing the source in *multiple* layers, and illustrate how source and channel layering schemes can be devised jointly with cooperative strategies.

IV. Distortion Exponent of the Half-duplex Relay Channel

In this section, we consider a half-duplex relay system, i.e., the relay terminal cannot receive and transmit simultaneously. We propose several joint source-channel coding and cooperation strategies for the half-duplex relay channel, and find their corresponding distortion exponents as a function of the bandwidth ratio.

A. Layered Source with Progressive Coding (LS)

In this strategy, we consider layered source compression followed by progressive transmission of the compressed layers over the channel. In layered source coding, each layer contains the refinement information for the preceding layers. By adjusting the compression and transmission rates, our goal is to provide more protection for the most important source layers, thereby improving the distortion exponent. Also, we will observe that cooperation can be used as a form of unequal error protection.

Suppose that the source encoder has n layers and the N channel uses are divided among these layers such that layer i is transmitted in t_iN channel uses, where $t_i \geq 0$, and $\sum_{i=1}^{N} t_i = 1$ (see Fig. 4-a). We will initially consider same transmission scheme for all layers. Let R_i bits per channel use be the channel transmission rate of the i-th layer. We denote the outage probability of layer i with the underlying transmission scheme as P_{out}^i and the distortion achieved when first i layers are successfully decoded as D_i^{LS}. Then using the successive refinability of the Gaussian source [29], for $i = 1, \ldots, n$, we have

$$D_i^{LS} = D\left(b\sum_{k=1}^{i} t_k R_k\right) = 2^{-b\sum_{k=1}^{i} t_k R_k},$$

with $D_0^{LS} = 1$. Note that, due to successive refinement source coding, layer $i+1, \ldots, n$ are useless when any of the preceding layers is missing. This forces us to assign transmission rates such that $P_{out}^i < P_{out}^{i+1}$, i.e., $R_i < R_{i+1}$. Then the expected distortion (ED) can be written as

$$\mathrm{ED}(\mathbf{R}, \mathbf{t}, \mathrm{SNR}) = \sum_{i=0}^{n} \left(P_{out}^{i+1} - P_{out}^{i}\right) \cdot D_i^{LS}, \quad (11)$$

where we define $P_{out}^0 = 0$, $P_{out}^{n+1} = 1$, and $\mathbf{R} = [R_1, \ldots, R_n]$, $\mathbf{t} = [t_1, \ldots, t_n]$. For any given average received SNR, we can write the minimum expected distortion problem as:

$$\begin{aligned} \min_{\mathbf{R},\mathbf{t}} \quad & \mathrm{ED}(\mathbf{R}, \mathbf{t}, \mathrm{SNR}) \\ \text{s.t.} \quad & \textstyle\sum_{i=1}^{n} t_i = 1, \\ & t_i \geq 0, \ \text{for } i = 1, \ldots, n \\ & 0 < R_1 < R_2 < \cdots < R_n. \end{aligned} \quad (12)$$

Recently [17] proposed an iterative algorithm to solve this optimization problem for the MIMO point-to-point channel. Although [17] shows the convergence of the proposed algorithm in small number of iterations by simulation, currently no linear-time algorithm is known to solve (12).

Our main focus here is the high SNR behavior, where we are able to obtain explicit results for the distortion exponent. As in the single rate scheme, for high SNR we have $R_i = r_i \log \mathrm{SNR}$, $i = 1, \ldots, n$. Let $d^*(r)$ be the optimal diversity gain corresponding to multiplexing gain of r for the underlying transmission protocol. Then in the high SNR regime, the expected distortion expression in (11) can be written as:

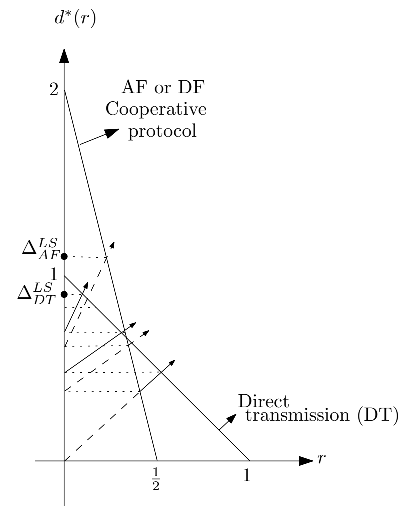

Fig. 5. A geometric interpretation of the LS strategy in the case of multiple source layers. Straight lines climb the DT curve, while the dashed ones climb the cooperation curve.

$$\begin{aligned} \mathrm{ED}(\mathbf{R}, \mathbf{t}, \mathrm{SNR}) &\doteq \sum_{i=0}^{n} \left[\mathrm{SNR}^{-d^*(r_{i+1})} - \mathrm{SNR}^{-d^*(r_i)}\right] \\ &\quad \cdot \mathrm{SNR}^{-b\sum_{k=1}^{i} t_k r_k} \\ &\doteq \sum_{i=0}^{n} \mathrm{SNR}^{-d^*(r_{i+1})} \mathrm{SNR}^{-b\sum_{k=1}^{i} t_k r_k} \quad (13) \end{aligned}$$

$$\doteq \mathrm{SNR}^{\max_{0\leq i\leq n}\left(-d^*(r_{i+1}) - b\sum_{k=1}^{i} t_k r_k\right)}, \quad (14)$$

where $d^*(r_{n+1}) = 0$, $0 \leq r_1 < r_2 < \cdots < r_n \leq 1$, and the last exponential equality is due to the fact that the summation is dominated by the slowest decay in the high SNR regime.

In [1], we showed that the optimal distortion exponent for LS strategy with finite number of layers can be written as a linear programming problem. Furthermore, it is shown that the optimal multiplexing gain allocation can be found as a solution of the following set of equations, that is, when all the terms in the summation in Eqn. (13) have equal exponents.

$$b t_n r_n = d^*(r_n), \quad (15)$$

$$d^*(r_n) + b t_{n-1} r_{n-1} = d^*(r_{n-1}), \quad (16)$$

$$\cdots$$

$$d^*(r_2) + b t_1 r_1 = d^*(r_1). \quad (17)$$

The above set of equations has an appealing interpretation on the DMT curve of the cooperative protocol that is used. As seen in Fig. 5, we have n straight lines ($i = 1, \ldots, n$). The i-th line has a slope of bt_i and intersect the y-axis at $d^*(r_{i+1})$, where $d^*(r_{i+1})$ is the y-ordinate of the intersection point of the trade-off curve with the $(i-1)$-th line. We assume $r_{n+1} = 1$, so $d^*(r_{n+1}) = 0$. This is similar to climbing up the DMT curve using an n-step staircase with varying step sizes. Notice that the y-ordinate of the topmost point gives us $d^*(r_1)$, the distortion exponent of n-layer LS, and the x-coordinates of the intersection points are the multiplexing gains of the

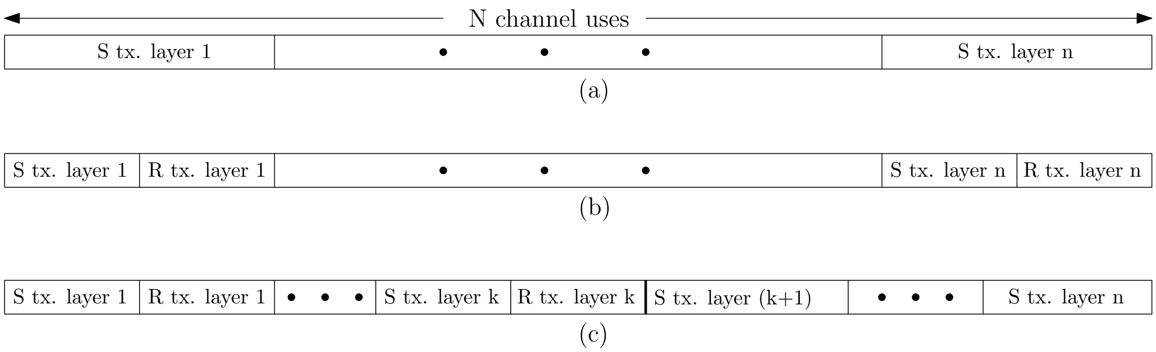

Fig. 6. Illustration of LS strategy with (a) only direct transmission, (b) only simple cooperative relaying, and (c) with both cooperative relaying and direct transmission.

corresponding layers. This interpretation allows us to observe that, increasing the number of layers improves the distortion exponent even though the total slope of the lines is equal to b.

In [1], we proved that, in the limit of infinite layers, for a transmission scheme with piecewise linear DMT curve, allocating the channel equally among the layers is optimal in terms of the distortion exponent, that is $t_i = 1/n$, for $i = 1, \dots, n$. This enabled us to obtain an explicit result for the distortion exponent of communication over MIMO channels in the limit of infinite layers, which serves as an upper bound for LS scheme with any finite number of layers. Since AF/DF relaying protocols have linear DMT curves, the same principles can be used to find the distortion exponent for LS, when all layers are transmitted using AF or DF. The following corollary gives a counterpart to Theorem 3.1 for infinite layer LS.

Corollary 4.1: LS with infinite layers results in the following distortion exponents for direct transmission and AF/DF cooperation protocols:

$$\Delta_{DT}^{LS} = 1 - e^{-b}, \tag{18}$$

$$\Delta_{AF}^{LS} = \Delta_{DF}^{LS} = 2(1 - e^{-b/4}). \tag{19}$$

Proof: Δ_{DT}^{LS} is given in [1]. Δ_{AF}^{LS} and Δ_{DF}^{LS} can be similarly derived using the analysis in [1] and the AF/DF DMT curve. ■

Comparing AF/DF and direct transmission (DT) DMT performances, we see that DT achieves a better diversity gain for high multiplexing gains. In general, for LS we can use any cooperation protocol or direct transmission for each layer within the channel uses allocated to it. This suggests that transmitting the first source layers -which are more essential from the end-to-end distortion perspective- with cooperation (hence better protection), while transmitting the rest directly (hence higher multiplexing gain), will improve the distortion exponent. This transmission strategy can be thought of providing unequal error protection through cooperation and is illustrated in Fig. 6-c. For comparison, LS with only DT and LS with only AF or DF is shown in Fig. 6-a and 6-b, respectively. The benefits of cooperating only on the more important source layers as a means to provide unequal error protection was also observed for a more practical system with finite SNR in [21], where we studied two-layer cooperative source-channel coding for a Gaussian source as well as real video using RCPC channel coding.

Visualizing this approach on Fig. 5, we observe that, we can achieve a higher distortion exponent, i.e., climb to a higher point, using direct transmission DMT for the last few layers. Once we pass the intersection point of DT and AF/DF DMT curves, using the cooperative tradeoff is more advantageous. This is equivalent to operating on the tradeoff curve that is obtained by picking the maximum diversity gain among the two protocols at each multiplexing gain, i.e., $d^*_{AF-DT}(r) = \max\{d^*_{DT}(r), d^*_{AF}(r)\}$. The portion of the layers that will be transmitted directly depends on the bandwidth ratio. When the bandwidth ratio is low, it is better to transmit all layers directly since the channel bandwidth is scarce and we would rather not use it for cooperation. On the other hand, if the bandwidth ratio is higher, we can transmit the lower layers with cooperation thus obtaining higher robustness for these important layers.

Corollary 4.2: Consider LS strategy where each layer is transmitted either with AF/DF relaying or directly as in Fig. 6-c. The distortion exponent in the limit of infinite layers is

$$\Delta_{AF-DT}^{LS} = \begin{cases} 1 - e^{-b} & \text{if} \quad b \leq \ln 3 \\ 2 - 4 \cdot 3^{-3/4} e^{-b/4} & \text{if} \quad b > \ln 3. \end{cases} \tag{20}$$

Proof: Proof is similar to Corollary 4.1 and is omitted. ■

Relating the distortion exponent of LS to the DMT curve of the transmission protocol as in Fig. 5 suggests that, further improvement in the distortion exponent is possible by using a cooperation protocol with a better DMT. For half-duplex relay systems, one of the best performing protocols in terms of DMT is dynamic decode-and-forward (DDF) proposed by Azarian et al. [6]. In the DDF protocol, the source terminal transmits for the entire time slot, while the relay listens until it can successfully decode the message. Once the relay can decode, it starts transmitting together with the source. The tradeoff for DDF protocol is given as [6]

$$d^*_{DDF}(r) = \begin{cases} 2(1-r) & \text{if} \quad 1/2 \geq r \geq 0 \\ (1-r)/r & \text{if} \quad 1 \geq r \geq 1/2. \end{cases} \tag{21}$$

Note that, diversity gain of DDF protocol dominates direct transmission, AF and DF at all multiplexing gain values, so it is best to transmit all layers using DDF. However, DDF requires more complex coding schemes such as incremental

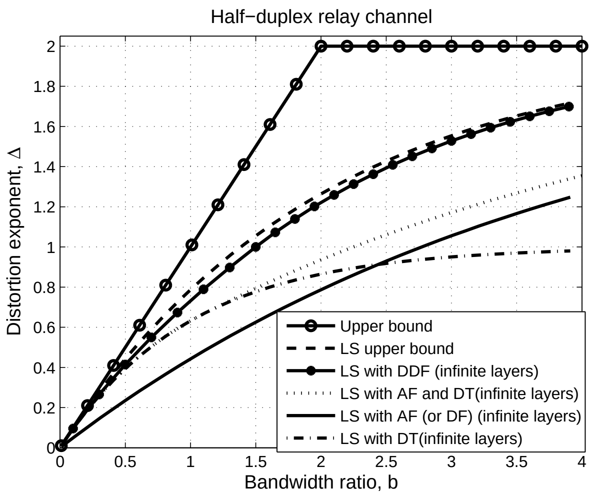

Fig. 7. Distortion exponent vs. bandwidth ratio for LS using different cooperation protocols.

redundancy codes for the source and space-time codes for the joint source-relay transmission.

Theorem 4.3: Consider a LS strategy that uses DDF protocol for all the layers. In the limit of infinite layers the distortion exponent is

$$\Delta_{DDF}^{LS} = \begin{cases} -1+\sqrt{1+2b} & \text{if} \quad 0 \leq b \leq 3/2 \\ 2-e^{3/4}e^{-b/2} & \text{if} \quad 3/2 \leq b. \end{cases} \tag{22}$$

Proof: In this case, the distortion exponent cannot be obtained directly from [1], since the tradeoff curve is not a piecewise linear function anymore. Appendix I provides the details. ■

In Fig. 7 we plot the distortion exponent vs. bandwidth ratio for LS coupled with various transmission strategies and the upper bound Δ^{UB} of Section III. We can observe the benefit of AF/DF type simple cooperation over DT at high bandwidth ratios. Although the distortion exponent of cooperative schemes converge to 2 as $b \to \infty$, while DT distortion exponent is limited to 1, the rate of convergence is very slow. LS with AF-DT as in Corollary 4.2 performs better than each strategy alone. As expected, DDF protocol has a considerable improvement over the other strategies, however, there is still a significant gap from the upper bound. We also included the distortion exponent upper bound for the LS scheme, by considering LS transmission over 2×1 MISO system. The gap between the DDF tradeoff curve and the DMT upper bound for multiplexing gains above $1/2$ reflects in the distortion exponent as well. We also note that LS distortion exponent upper bound is still far away from Δ^{UB} of Section III.

B. Hybrid Digital-Analog Transmission with Layered Source (HLS)

In general, digital transmission suffers from the threshold effect, that is, the system performance degrades drastically when the instantaneous channel capacity falls below the target rate. LS strategy introduces source scalability and thus adaptivity to the channel variations to partially overcome the threshold effect. Another technique in the literature that successfully mitigates the threshold effect is hybrid digital-analog transmission. Mittal and Phamdo [32] introduced hybrid digital-analog transmission to improve the end-to-end distortion performance when transmitting over an unknown noise channel or broadcasting to users with different noise variances. For practical construction of hybrid schemes, see [33]- [35] and references therein.

In [12], we considered transmitting source samples in a pure analog fashion, i.e., simply by power scaling. Although this is optimal for a point-to-point single antenna system for $b \geq 1$ [13], when the relay is present, pure analog transmission cannot utilize the additional diversity introduced. Hybrid digital-analog transmission has the potential to both mitigate threshold effect and utilize the inherent diversity in the system, and hence improve the distortion exponent as well [10].

For $b \leq 1$, we use the hybrid scheme of [10] which is based on dividing the source samples into two groups, where one group is coded, the other one is simply scaled, and the two groups are superimposed for transmission. This strategy achieves $\Delta = b$, which is equal to Δ^{UB} for $b \leq 1$ and thus, this hybrid scheme for direct transmission is optimal for the relay scenario as well.

For $b > 1$, we propose hybrid digital-analog transmission with layered source (HLS) strategy that combines LS with analog transmission [1]. This can be considered as a combination of the hybrid digital-analog scheme of [10] and our LS strategy. In HLS, we divide N channel uses into two portions. The first portion, composed of $N-K$ channel uses, is reserved for layered source coding similar to LS, where the second portion, composed of K channel uses, is used for uncoded transmission as illustrated in Fig. 4-b. Just as in the LS, we assume that the source samples are compressed into n layers where layer k contains the successive refinement information for layers $1, \ldots, k-1$. These source layers can be transmitted using direct transmission or cooperation. Let $\bar{\mathbf{s}} \in \mathbb{C}^K$ be the reconstruction of the source vector $\mathbf{s} = [s_1, \ldots, s_K]$ upon successful reception of all the LS layers. We denote the reconstruction error as $\mathbf{e} \in \mathbb{C}^K$ where $\mathbf{e} = \mathbf{s} - \bar{\mathbf{s}}$. We transmit this error during the second portion of K channel uses where each component of the error vector is transmitted without coding in an analog fashion. We simply scale the error vector such that the transmit power constraint is satisfied. This analog transmission is directly from the source terminal to the destination without cooperation. This is justified since relaying does not improve the distortion exponent behavior of analog transmission as discussed above.

The destination first attempts to decode all the digitally transmitted layers, and in case of successful reception of all the layers, it forms the estimate $\bar{\mathbf{s}} + \tilde{\mathbf{e}}$, where $\tilde{\mathbf{e}}$ is the linear MMSE estimate of $\mathbf{e}$ based on the received signal during K channel uses reserved for analog transmission. This analog portion is neglected unless all digitally transmitted layers can be successfully decoded at the destination. Similar to Section IV-A, we can find the HLS distortion exponent as stated in the following theorem.

Theorem 4.4: For HLS where each layer is transmitted

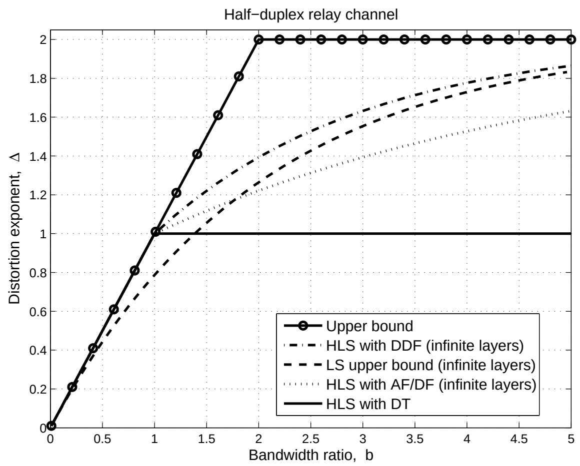

Fig. 8. Distortion exponent vs. bandwidth ratio of HLS strategy with different cooperation protocols.

by cooperation using AF/DF or DDF protocols, respectively, and the error corresponding to the last layer is transmitted uncoded directly from the source terminal to the destination, the distortion exponent in the limit of infinite layers for $b > 1$ is

$$\Delta_{AF}^{HLS} = \Delta_{DF}^{HLS} = 2 - e^{-(b-1)/4}, \quad (23)$$
$$\Delta_{DDF}^{HLS} = 2 - e^{-(b-1)/2}. \quad (24)$$

Proof: See Appendix II. ■

The proof of Theorem 4.4 suggests that the distortion exponent of HLS can be found using the DMT curve of the transmission protocol in a manner similar to LS. However, with HLS while climbing on the DMT curve, the first line passes through $(0, 1)$ instead of the origin, and the total slope is reduced to $b - 1$.

In Figure 8, we illustrate the distortion exponents of various cooperation protocols with HLS. We also included the LS upper bound of Section IV-A for comparison. Note that the HLS upper bound is achievable by a half-duplex relay, since HLS strategy utilizes the DMT curve for diversity gains greater than 1 and DDF is DMT optimal in this range. The figure suggests that, for small bandwidth ratios, HLS improves the distortion exponent compared to the LS upper bound, even when used with a simple cooperation scheme, such as AF or DF. For higher bandwidth ratios, we need to utilize a more advanced cooperation protocol such as DDF to operate on a better DMT curve. However, if the transmission protocol is fixed, HLS always achieves a higher distortion exponent than LS with the same number of source layers, while the improvement decays to zero as bandwidth ratio increases.

C. Broadcast Strategy with Layered Source (BS)

Broadcast approach introduced in [36], [37] proposes to consider the fading channel with unknown state information at the transmitter as a broadcast channel, where each receiving user's channel corresponds to a different realization of the fading. Based on this model, LS strategy introduced above would correspond to transmitting to users through time-sharing. However, as proved in [38], superposition coding performs better than time-sharing for broadcast channels. Assuming successive decoding at the receiver, information intended for a bad channel state can also be received when the channel quality is better. This information can then be decoded and subtracted from the received signal, enabling the reception of additional information depending on the channel state. Hence, the better the fading channel state, the more information can be received at the destination terminal. Broadcast strategy for MIMO channels is considered in [39] and it is shown that it improves the throughput of the system. In [12], [13], we combined this broadcast idea with layered source coding calling it *broadcast strategy with layered source* (BS). We showed that BS strategy significantly improves the distortion exponent. In fact, it can achieve the optimal distortion exponent for MISO/SIMO systems for all bandwidth ratios, while it is optimal for general block fading MIMO systems for high bandwidth ratios [1].

In this section, we consider BS strategy for cooperative systems, where superimposed layers will be transmitted to the destination using various cooperation protocols. As discussed in Section IV-A, in LS strategy we have time-division among transmission of source layers, and each layer can be transmitted using a different cooperation protocol, with the goal of climbing up the tradeoff curve as high as possible. On the other hand, as we will observe in this section, in the BS case, combining superimposed transmission of source layers with a cooperation protocol, and allocating the transmit power and the transmission rates among the layers cannot be done in a straightforward manner.

We again assume that the source is compressed in n layers where layer i is composed of the successive refinement information for layers $1, \ldots, i-1$ for $i = 2, \ldots, n$. The source layers are channel coded separately and transmitted simultaneously, while coding rates, allocated powers and the protocol used for each layer (to cooperate or to transmit directly, what cooperation protocol to use) are optimized jointly considering the overall system performance. See Fig. 4-c for an illustration of the general BS strategy. The decoder at the destination terminal is assumed to perform successive decoding, where layers are decoded in order starting from the first layer, and the decoded codeword for each layer is subtracted from the received signal. When the decoder cannot decode a layer, it declares outage for that layer and all the following ones.

We define the set of network channel states $\mathbf{h} = (h_1, h_2, h_3)$ (see Fig. 1) that result in outage for layer i at the destination after decoding and subtracting the previous layers as $\mathcal{O}_i^d$. Then the overall set of channel states for which layer i is in outage at the destination can be written as $\bar{\mathcal{O}}_i^d = \bigcup_{k=1}^{i} \mathcal{O}_k^d$. We define the corresponding outage probabilities as follows.

$$P_{out}^{d,i} = Pr\{\mathbf{h} : \mathbf{h} \in \mathcal{O}_i^d\}, \quad (25)$$
$$\bar{P}_{out}^{d,i} = Pr\{\mathbf{h} : \mathbf{h} \in \bar{\mathcal{O}}_i^d\}. \quad (26)$$

These outage probabilities depend on the rate and power allocation among the layers as well as the cooperation scheme used. Let $\mathbf{R} = [R_1, \ldots, R_n]^T$ be the channel rate vector. If we define the average distortion achieved by receiving the first k layers as D_i^{BS}, we can write

$$D_i^{BS} = D\left(b\sum_{k=1}^{i} R_k\right), \quad (27)$$

where we use the successive refinability of the source. Then we can write the expected distortion for n layer BS strategy using successive decoding as

$$\mathrm{ED}(\mathbf{R}, \mathrm{SNR}) = \sum_{i=0}^{n}(\bar{P}_{out}^{d,i+1} - \bar{P}_{out}^{d,i})D_i^{BS}, \quad (28)$$

with $\bar{P}_{out}^{d,0} = 0$, $\bar{P}_{out}^{d,n+1} = 1$, and $D_0^{BS} = 1$. Similar to Section IV-A, we can argue that the rate of each layer should scale as $O(\log \mathrm{SNR})$ with increasing signal-to-noise ratio in order to have vanishing expected distortion. Let the multiplexing gain vector be $\mathbf{r} = [r_1, \ldots, r_n]^T$. We have $\mathbf{R} = \mathbf{r}\log\mathrm{SNR}$, and the high SNR approximation for expected distortion can be obtained as

$$\mathrm{ED}(\mathbf{r}, \mathrm{SNR}) \doteq \sum_{i=0}^{n} \bar{P}_{out}^{d,i+1}\mathrm{SNR}^{-b\sum_{j=1}^{i} r_j}. \quad (29)$$

We define the *successive decoding diversity gain* of each layer at the destination based on the outage probability of successive decoding as $d_{sd}(r_i)$, i.e., $\bar{P}_{out}^{d,i} \doteq \mathrm{SNR}^{-d_{sd}(r_i)}$. Then we can rewrite the expected distortion as below.

$$\mathrm{ED}(\mathbf{r}, \mathrm{SNR}) \doteq \sum_{i=0}^{n} \mathrm{SNR}^{-d_{sd}(r_{i+1})}\mathrm{SNR}^{-b\sum_{j=1}^{i} r_j}. \quad (30)$$

Based on this expression, since the slowest decay rate will dominate the expected distortion, we can find the distortion exponent of BS strategy for the given power and multiplexing gain allocation and the underlying transmission policy as

$$\Delta = \min_{0\le i\le n}\left\{d_{sd}(r_{i+1}) + b\sum_{j=1}^{i} r_j\right\}. \quad (31)$$

Recall that in LS, the DMT curve of the underlying transmission scheme is sufficient to determine the best achievable distortion exponent. Similarly, when BS is used, once we are given the successive decoding diversity-multiplexing gain tradeoff curve of the underlying transmission scheme obtained by the best power allocation among layers, we can find the corresponding highest distortion exponent achieved by BS strategy. Let $d^*(r)$ be the original DMT curve for the given transmission scheme obtained when transmitting a single codeword with full power. We can bound the successive decoding diversity gain of the i-th layer as $d_{sd}(r_i) \le d^*(r_1 + \ldots + r_i)$, since successful transmission of the i-th layer with BS requires successful transmission of all the preceding layers, which is equivalent to transmitting a total rate of $(r_1 + \ldots + r_i)\log\mathrm{SNR}$.

In [1], we proposed a power allocation scheme for MISO/SIMO systems such that $d_{sd}(r_i) = d^*(r_1 + \ldots + r_i)$ for all $i = 1, \ldots, n$. This condition suggests that all layers can simultaneously operate on the optimal DMT curve for all multiplexing gain allocations. In an independent study, Diggavi and Tse also consider the problem of transmitting multiple streams over a MIMO channel with different multiplexing and diversity gains [40]. They coin the term *successive refinability of the diversity-multiplexing tradeoff curve* for the cases when the reliability of receiving layers up to k is equal to the reliability one would achieve sending only these layers. Hence, having a successively refinable DMT curve is equivalent to successive decoding diversity gain lying on the DMT curve.

We next investigate various cooperation protocols and power allocation among layers to study the successive decoding diversity gain. We will show that, a power allocation rule similar to the one used for the SIMO/MISO systems in [1] will successively refine the diversity-multiplexing tradeoff curve of some cooperation protocols we study, while extension to other protocols is not straightforward.

1) Amplify-and-Forward (AF): We first consider the AF cooperation protocol for each of the transmitted layers. In BS with AF, N uses of each channel realization are divided into two equal portions. While the source transmits a superposition of the channel codewords of n layers with an appropriate power allocation in the first half, the relay simply amplifies and forwards its received signal within its power constraint in the second half. Let the received signals at the relay and the destination at time instant i be $y_r[i]$ and $y_d[i]$, respectively. We illustrate the BS strategy with AF cooperation protocol in Fig. 9-b. We have,

$$y_r[i] = h_2\sum_{j=1}^{n}\sqrt{\mathrm{SNR}_j}x_{s,j}[i] + z_r[i], \text{ for } i = 1, \ldots, N/2$$

$$y_d[i] = h_1\sum_{j=1}^{n}\sqrt{\mathrm{SNR}_j}x_{s,j}[i] + z_d[i], \text{ for } i = 1, \ldots, N/2$$

$$y_d[i] = h_3\beta y_r[i - N/2] + z_d[i], \text{ for } i = N/2+1, \ldots, N$$

where $z_r[i]$ and $z_d[i]$ are independent additive circularly symmetric complex Gaussian noise terms with variance $1/2$ in each dimension, $\sqrt{\mathrm{SNR}_j}x_{s,j}[i]$ is the source signal transmitted at time instant i, corresponding to the j-th source layer, and $\beta y_r[i - N/2]$ is the relay signal transmitted at time i. Due to power constraint at the source, we have $\frac{1}{N}\sum_{i=1}^{N} x_{s,j}[i] = 1$, while $\sum_{j=1}^{n}\mathrm{SNR}_j \le \mathrm{SNR}$. For the relay to satisfy its power constraint, we need

$$\beta = \sqrt{\frac{\mathrm{SNR}}{|h_2|^2\mathrm{SNR} + 1}}. \quad (32)$$

Theorem 4.5: The optimal distortion exponent of AF cooperation protocol with BS, in terms of the number of layers n, and bandwidth ratio b, is

$$\Delta_{AF,n}^{BS} = 2 - \frac{8(4-b)}{16 + (32 - b^2)(\frac{b}{4})^{n-1}}. \quad (33)$$

This is an increasing function of n, and in the limit of infinite layers, we obtain

$$\Delta_{AF}^{BS} = \lim_{n\to\infty}\Delta_{AF,n}^{BS} = \begin{cases} b/2 & \text{if } 0 < b < 4 \\ 2 & \text{if } 4 \le b. \end{cases} \quad (34)$$

Proof: We define the total power allocated to layers $k, k+1, \ldots, n$ as $\overline{\mathrm{SNR}}_k \triangleq \sum_{i=k}^{n}\mathrm{SNR}_i$ for $k = 1, \ldots, n$. We have

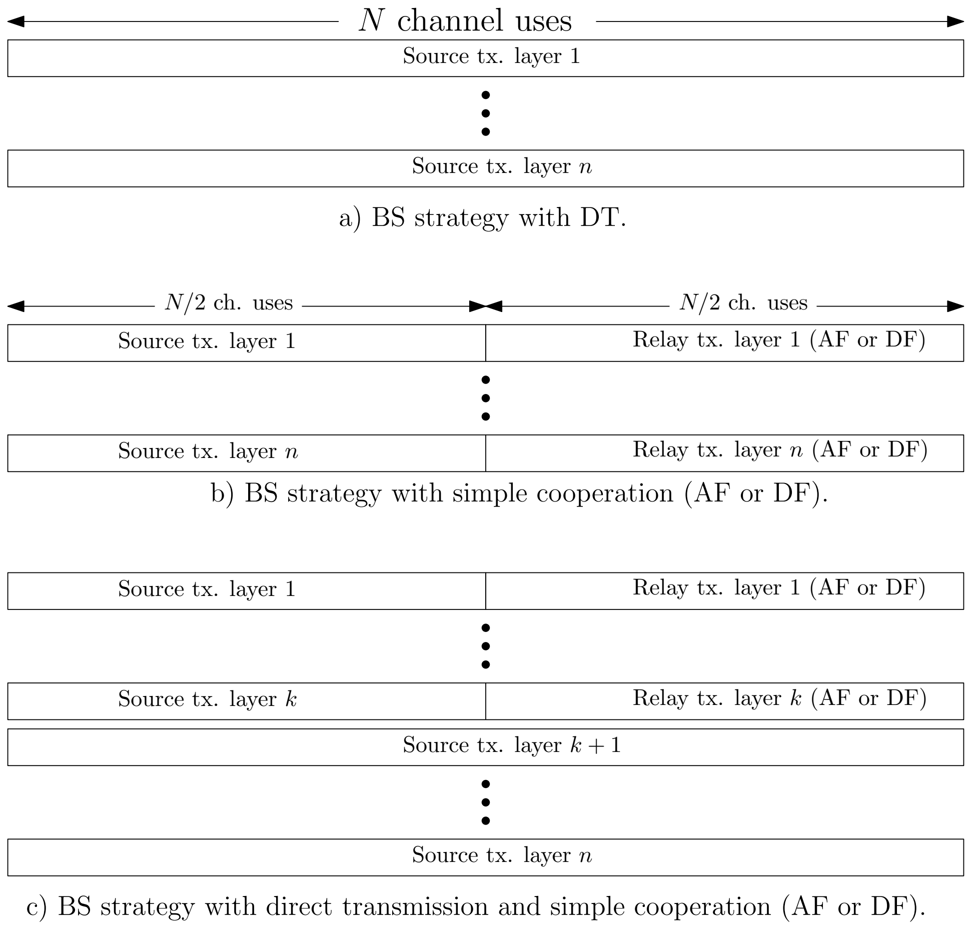

Fig. 9. BS strategy coupled with direct transmission or simple cooperative schemes.

$\overline{\mathrm{SNR}}_1 = \mathrm{SNR}$. In order to ensure successive refinement of the AF DMT curve we use an exponential power allocation strategy similar to the one used in [12], [13]. For $k = 2, \dots, n$, and multiplexing gain vector $\mathbf{r} = [r_1, \dots, r_n]$ let

$$\overline{\mathrm{SNR}}_k = \mathrm{SNR}^{1-2r_1-\cdots-2r_{k-1}-\epsilon_{k-1}}, \tag{35}$$

for some $0 < \epsilon_1 < \cdots < \epsilon_{n-1}$. Then we have

$$\mathrm{SNR}_i = \mathrm{SNR}^{1-(2r_1+\cdots+2r_{i-1}+\epsilon_{i-1})} - \mathrm{SNR}^{1-(2r_1+\cdots+2r_i+\epsilon_i)}, \tag{36}$$

for $i = 1, \dots, n-1$, and

$$\mathrm{SNR}_n = \mathrm{SNR}^{1-(2r_1+\cdots+2r_{n-1}+\epsilon_{n-1})}.$$

We show in Appendix III that, with $\epsilon_{n-1} \to 0$, the successive decoding diversity gain of the AF protocol becomes $d_{sd}^{AF}(r_i) = d_{AF}^*(r_1 + \cdots + r_i) = 2(1 - 2r_1 - \cdots - 2r_i)$, for $i = 1, \dots, n$. Hence, AF protocol is successively refinable in the DMT sense.

Once we have the successive decoding diversity gain for AF protocol, we can rewrite the distortion exponent from Eqn. (31).

$$\Delta = \min_{0 \le i \le n} \left\{ 2 - 4(r_1 + \cdots + r_{i+1}) + b(r_1 + \cdots + r_i) \right\},$$

where we define $r_{n+1} = 1 - (r_1 + \cdots + r_n)$. The optimal distortion exponent for this transmission scheme is obtained when all the terms in the above minimization are equal. It is easy to check that, this is achieved by the following multiplexing gain allocation, leading to eqn. (33).

$$r_1 = \frac{2(4-b)}{16 + (32 - b^2)\left(\frac{b}{4}\right)^{n-1}},$$

$$r_i = \left(\frac{b}{4}\right)^{i-1} r_1, \text{ for } i = 2, \dots, n.$$

■

2) Decode-and-Forward (DF): Next we consider decode-and-forward (DF) as our cooperation protocol. Recall that the DF protocol of [5] achieves the same DMT curve as of AF. Although in the LS strategy, this equivalence of the DMT curves would suggest an equality of the achievable distortion exponents, in the BS strategy, we cannot claim this directly, since we need the equivalence of the successive decoding DMT curves. In our analysis, we will generalize the DF protocol in [5] to unequal division of the channel between the source and the relay transmissions, and find the conditions under which the resulting DMT is successively refinable. In the generalized DF protocol we consider, the source transmits for the first αN channel uses ($\alpha \ge 1/2$), and for the rest of the $(1-\alpha)N$ channel uses only the relay terminal transmits if it can decode the source information. We assume that the relay reencodes the original message using an independent Gaussian codebook. Note that α does not depend on the instantaneous fading state, so this can be considered as a static protocol.

Lemma 4.6: The DMT for the generalized DF protocol for fixed α is characterized by

$$d^*_{gDF}(r) = \begin{cases} \min\left(2 - \frac{r}{1-\alpha}, 2 - \frac{2r}{\alpha}\right) & \text{if} \quad r \leq 1-\alpha \\ \min\left(2 - \frac{2r}{\alpha}, \frac{1-r}{\alpha}\right) & \text{if} \quad 1-\alpha \leq r \leq \alpha \\ 0 & \text{if} \quad r > \alpha. \end{cases} \tag{37}$$

Proof: Proof can be found in Appendix IV. ∎

Note that, for $\alpha = 1/2$ we get the classical DF protocol with DMT $d^*(r) = 2(1-2r)$ for $r \in [0, 1/2]$. On the other hand, for $\alpha = 2/3$, we have $d^*(r) = 2 - 3r$ for $r \in [0, 2/3]$ which achieves the best diversity gain over all generalized DF schemes for multiplexing gains less than $1/3$. For each multiplexing gain $r \geq 1/3$, the diversity gain is maximized for a different channel allocation α.

Now we consider using BS strategy coupled with the generalized DF protocol. While the source transmits a superposition of the channel codewords for the n source layers in the first αN channel uses, the relay terminal tries to decode as many layers as possible, and retransmits those layers in the rest of the $(1-\alpha)N$ channel uses. Note that, while the diversity gain is maximized for a different channel allocation for each $r \geq 1/3$, we cannot use a different α for each source layer due to the half-duplex constraint on the relay. This constrains us to a fixed channel allocation for transmission of all the layers.

We assume successive decoding at both the relay and the destination. Let y_r and y_d be the received signals at the relay and the destination, respectively. Assuming m source layers can be successfully decoded at the relay, the channel model for generalized DF protocol with BS strategy can be written as below.

$$y_r[i] = h_2 \sum_{k=1}^{n} \sqrt{\mathrm{SNR}^s_k} x_{s,k}[i] + z_r[i], \tag{38}$$

$$y_d[i] = h_1 \sum_{k=1}^{n} \sqrt{\mathrm{SNR}^s_k} x_{s,k}[i] + z_d[i], \tag{39}$$

for $i = 1, \ldots, \alpha N$, and

$$y_d[i] = h_3 \sum_{k=1}^{m} \sqrt{\mathrm{SNR}^r_k} x_{r,k}[i] + z_d[i], \tag{40}$$

for $i = \alpha N + 1, \ldots, N$, where $z_r[i]$ and $z_d[i]$ are i.i.d. additive white Gaussian noise terms as before, and $\sqrt{\mathrm{SNR}^s_i} x_{s,k}[i]$, $\sqrt{\mathrm{SNR}^r_k} x_{r,k}[i]$ are the source and the relay signals, respectively, corresponding to the k-th source layer transmitted at time i. Due to power constraints at the source and the relay, we have $\frac{1}{N}\sum_{i=1}^{N} x_{s,k}[i] = 1$ and $\frac{1}{N}\sum_{i=1}^{N} x_{r,k}[i] = 1$ for all k, while $\sum_{k=1}^{n} \mathrm{SNR}^s_k \leq \mathrm{SNR}$ and $\sum_{k=1}^{n} \mathrm{SNR}^r_k \leq \mathrm{SNR}$.

Lemma 4.7: Consider the generalized DF protocol coupled with BS strategy with n source layers as explained above. Let $\mathbf{r} = [r_1, \ldots, r_n]$ be the multiplexing rate vector. We define $\overline{\mathrm{SNR}}^s_k = \sum_{i=k}^{n} \mathrm{SNR}^s_i$ and $\overline{\mathrm{SNR}}^r_k = \sum_{i=k}^{n} \mathrm{SNR}^r_i$. For the power allocation at the source terminal

$$\overline{\mathrm{SNR}}^s_k = \mathrm{SNR}^{1 - \frac{1}{\alpha}(r_1 + \cdots + r_{k-1} - \epsilon_{k-1})}, \tag{41}$$

and the power allocation at the relay terminal

$$\overline{\mathrm{SNR}}^r_k = \mathrm{SNR}^{1 - \frac{1}{1-\alpha}(r_1 + \cdots + r_{k-1} - \epsilon_{k-1})}, \tag{42}$$

we obtain the successive decoding diversity gain $d_{sd}(r_i)$ for layer i as

$$d_{sd}(r_i) = \min\left\{ 2\left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_i)\right)^+, \left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_i)\right)^+ + \left(1 - \frac{1}{1-\alpha}(r_1 + \cdots + r_{i-1})\right)^+ \right\}, \tag{43}$$

where we define $x^+ \triangleq \max\{x, 0\}$.

Proof: Proof can be found in Appendix V. ∎

The following corollary which directly follows from Lemma 4.6 and Lemma 4.7, states that the classical DF protocol ($\alpha = 1/2$) indeed achieves the same distortion exponent as AF.

Corollary 4.8: Classical DF protocol with $\alpha = 1/2$ is successively refinable in terms of the DMT curve. Thus, with the same number of source layers, it can achieve the same distortion exponent as AF. Therefore $\Delta^{BS}_{DF,n} = \Delta^{BS}_{AF,n}$ and $\Delta^{BS}_{DF} = \Delta^{BS}_{AF}$, where $\Delta^{BS}_{AF,n}$ and Δ^{BS}_{AF} are given by (33)-(34).

Corollary 4.9: Consider the generalized DF protocol with $\alpha = 2/3$. This cooperation protocol is successively refinable in terms of the DMT curve if

$$r_k \geq r_1 + \cdots + r_{k-1}, \tag{44}$$

for all $k = 1, \ldots, n$.

Proof: Under condition (44), using Lemma 4.7 we can show that the successive decoding diversity gain of DF protocol with $\alpha = 2/3$ is

$$d^*_{sd}(r_k) = 2 - 3(r_1 + \cdots + r_k), \tag{45}$$

if $r_1 + \cdots + r_k \leq 2/3$ and 0 elsewhere. By Lemma 4.6, this is equal to $d^*(r_1 + \cdots + r_k)$, where $d^*(r)$ is the diversity gain of the generalized DF protocol with $\alpha = 2/3$. ∎

This example shows that, successive refinability of a DMT curve may not hold for all cooperation protocols, under all multiplexing gain allocations. In [41], authors prove that the DMT curve of two parallel fading channels is not successively refinable. Although we do not give a formal proof, we conjecture that generalized DF is not successively refinable unless $\alpha = 1/2$. This suggests that even though generalized DF for $\alpha = 2/3$ has higher DMT than $\alpha = 1/2$, it is not clear whether there exists a power allocation resulting in successive refinability of the DMT, or in general, whether a power allocation results in a successive decoding diversity gain that is higher than the DMT with $\alpha = 1/2$. We will revisit this observation later when we discuss static cooperation protocols in conjunction with BS.

3) Direct transmission combined with AF/DF relaying: Similar to Section IV-A sending important (the first few) source layers using cooperation and the rest directly may improve the distortion exponent further since direct transmission DMT is better than AF/DF cooperation for high multiplexing gains. For BS, rather than having separate time slots for cooperation and direct transmission as we did in Section IV-A for LS, we need to superimpose multiple layers corresponding to cooperation and direct transmission while ensuring good successive decoding diversity for each layer. For this, we consider a combination of direct transmission and DF with $\alpha = 1/2$ cooperation, where layers with low multiplexing and high diversity gain (important layers) are transmitted using DF while the rest are transmitted directly as shown in Fig. 9-c. The overall diversity gain of this protocol is characterized by $d^*_{DF-DT}(r) = max(1-r, 2-4r)$ for a multiplexing gain of $r \in [0,1]$. Similar results can be obtained by combining AF with DT as well.

Theorem 4.10: The optimal distortion exponent of the above DF-DT cooperation protocol with BS strategy consisting of n source layers, where we transmit the first $k = \beta n$ layers by cooperation and the following $(n-k)$ layers directly, is given by

$$\Delta^{BS}_{DF-DT,n} = \frac{4 - b + 2(b-1)\left(\frac{b}{4^{\beta}}\right)^n - (b+2)b^{(1-\beta)n}}{\frac{4}{b} - 1 + (b-1)\left(\frac{b}{4^{\beta}}\right)^n - 3b^{(1-\beta)n}}. \quad (46)$$

The distortion exponent in the limit of infinite layers, i.e., $n, k \to \infty$, is given as

$$\Delta^{BS}_{DF-DT} = \begin{cases} b & \text{if} \quad 0 < b < 1 \\ \frac{b+2}{3} & \text{if} \quad 1 \leq b < 4 \\ 2 & \text{if} \quad 4 \leq b. \end{cases} \quad (47)$$

Proof: Proof is given in Appendix VI. The proof uses the successive refinability of the DMT for half-duplex DF-DT scheme which is shown in Appendix VII. ∎

4) Discussion of the results: Comparison of the distortion exponents achieved by BS with different transmission schemes can be seen in Figure 10. This figure also contains HLS with DDF, the best non-broadcast type strategy we have studied. We note that HLS with DDF slightly outperforms BS with DF-DT for certain bandwidth ratios, while BS with DF-DT provides optimal distortion exponent for $b \leq 1$ and $b \geq 4$.

A cooperation protocol with a better diversity-multiplexing tradeoff has the potential to result in a higher distortion exponent using BS strategy. However, one may not always be able to efficiently combine BS with the cooperation protocol under consideration. Also, if we focus on successive decoding, we need the underlying transmission scheme not only to have a better DMT curve, but also to have a better successive decoding diversity gain. Ideally, we need a successively refinable DMT.

For example, consider transmitting a superposition of multiple layers through the DDF protocol [6]. Recall that, in DDF protocol the relay tries to decode the base layer, and when it succeeds, DDF suggests that it should reencode and start transmitting the base layer to the destination. However, since the relay cannot receive while transmitting, it cannot decode any further layers. Thus, it is not clear how to effectively combine a BS type strategy with a dynamic cooperation protocol such as DDF and what the optimal number of layers for the relay to decode would be. This urges us to concentrate only on static cooperation protocols with the same channel allocation over the layers.

The static protocol described in [42] outperforms all the other known static cooperation protocols in terms of DMT. This protocol combines generalized DF with $\alpha = 2/3$ with the source terminal transmitting additional information while the relay is forwarding. However, we have seen in Corollary 4.8 that even generalized DF with $\alpha = 2/3$ does not necessarily have a successively refinable DMT. We conclude that extending BS strategy to more complex protocols requires more investigation, and it is not clear a priori how much additional gain one can get in the distortion exponent by utilizing these extensions.

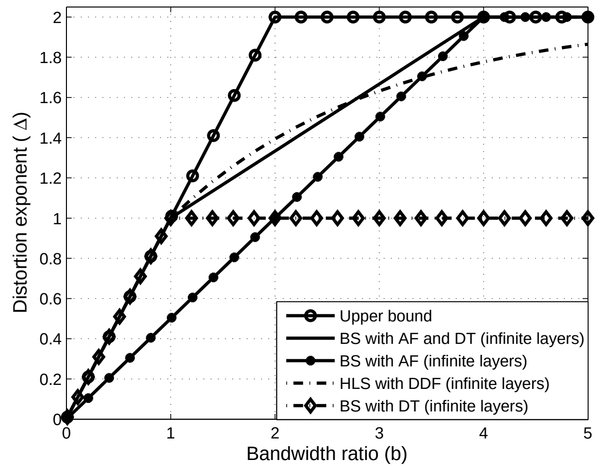

Fig. 10. Distortion exponent vs. bandwidth ratio of BS strategy using different cooperative protocols. For comparison, the full-duplex upper bound, and HLS with DDF are also shown.

V. Distortion Exponent with Full Duplex Relay

In this section we consider a full-duplex relay terminal and provide achievable distortion exponent results. Based on the analysis in Section IV, it is straightforward to extend the LS and HLS strategies to the full-duplex relay scenario, to obtain the following results.

Lemma 5.1: The distortion exponent of LS for a full-duplex relay in the limit of infinite layers is

$$\Delta^{LS}_{fd-DF} = 2(1 - e^{-b/2}), \quad (48)$$

while the distortion exponent of HLS in the limit of infinite layers is

$$\Delta^{HLS}_{fd-DF} = 2 - e^{-(b-1)/2}. \quad (49)$$

Proof: For a full-duplex cooperative system, decode-and-forward (DF) protocol is optimal in terms of diversity-multiplexing tradeoff, that is it achieves the same DMT curve of the 2×1 (or 1×2) MIMO system [7]. Since LS and HLS distortion exponents are completely determined by the DMT curve, their optimal distortion exponents will be exactly the same as the distortion exponent of a 2×1 MIMO channel with the corresponding source-channel coding strategy [1]. ∎

For BS strategy, as argued in Section IV-C, knowing the DMT is not sufficient, and we need to determine the successive decoding DMT to compute the distortion exponent. The proof of the following theorem shows that BS with full-duplex DF has a successively refinable DMT.

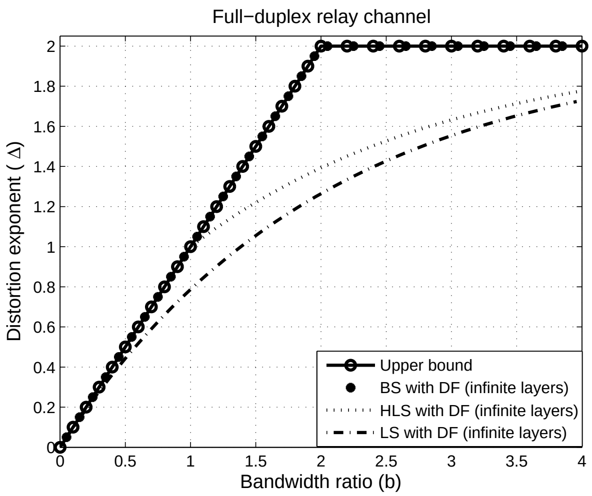

Fig. 11. Distortion exponent vs. bandwidth ratio for the full-duplex cooperative relay channel.

Theorem 5.2: The distortion exponent of full-duplex DF with n-layer BS is given as

$$\Delta^{BS}_{fd-DF,n} = 2\left(1 - \frac{1-b/2}{1-(b/2)^{n+1}}\right). \tag{50}$$

Proof: Proof is given in Appendix VIII. The proof uses successive decoding diversity of full-duplex DF which is found in Appendix IX. ■

In the limit of infinite layers, $\Delta^{BS}_{fd-DF,n}$ approaches the upper bound in Theorem 3.2, and we have the following corollary.

Corollary 5.3: BS strategy coupled with full-duplex DF relaying has optimal distortion exponent in the limit of infinite source coding layers. We have

$$\Delta^{BS}_{fd} = \lim_{n\to\infty} \Delta^{BS}_{fd-DF,n} = \begin{cases} b & \text{if} \quad b < 2, \\ 2 & \text{if} \quad b \geq 2. \end{cases} \tag{51}$$

Fig. 11 summarizes the distortion exponents of BS, HLS and LS all employing a full duplex DF cooperation protocol.

VI. Correlated Side Information

The original motivation behind user cooperation is to improve the reliability of wireless uplink channel [3], [4]. However, when source transmission is considered, the availability of correlated side information at the destination and/or at the relay terminals should be taken into account since it has the potential to improve the end-to-end distortion performance [45]. This may be especially relevant for sensor applications where user cooperation has been proposed as an effective tool to improve the sensor network reliability.

Under the assumption of correlated side-information at the relay terminal, source-channel separation theorem fails even when there is no fading. We investigate joint source-channel and relaying strategies for lossy transmission over an AWGN channel in [46] and for lossless transmission over discrete memoryless channel in [47], however our focus here is on the quasi-static fading model, and the effect of side information on the achievable distortion exponent. We next argue that the increase in distortion exponent with the introduction of the relay terminal is mainly due to the increase in channel diversity. We show that the availability of a correlated side information at the destination and/or the relay would not further improve either the upper bound or the achievable distortion exponents of the schemes we have considered.

Suppose that the destination and the relay have access to correlated side information sequences S_1 and S_2, respectively. We assume that $S_i = S + W_i$, $i = 1, 2$, where W_i is a zero-mean complex Gaussian with variance σ_i^2, and is independent of S and each other. No side information is equivalent to having $\sigma_i^2 \to \infty$. Suppose both S_1 and S_2 are made available to all three terminals. Note that, this can only reduce the achievable expected distortion, hence increase the distortion exponent. Under this assumption, we can now use conditional distortion-rate function $D_{S|S_1,S_2}(R)$ of the source S given (S_1, S_2), which for the Gaussian source and side information, is equal to

$$D_{S|S_1,S_2}(R) = \left(\frac{\tilde{\sigma}^2}{1+\tilde{\sigma}^2}\right) 2^{-R}, \tag{52}$$

where $\tilde{\sigma}^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2+\sigma_2^2}$.

We observe from (52) that, availability of side information does not change the exponential dependence of the rate-distortion function on R, it only reduces the effective variance. However, constant scaling in the distortion-rate function does not change the distortion exponent. Therefore, providing the side information (S_1, S_2) to all three terminals does not improve the upper bound or the achievable distortion exponents of the schemes we have considered. We can conclude that side-information does not help in increasing the distortion exponent of the strategies considered in this paper. However, when channel SNR is finite, side information aware joint source-channel coding techniques have the potential to reduce the expected distortion [46].

VII. Conclusion

In this work, we consider the problem of joint source channel coding over slowly fading cooperative relay systems. Our focus is on the high SNR behavior of the expected distortion characterized by the distortion exponent. Even though single layer source coding followed by cooperative transmission improves the achievable distortion exponent for high bandwidth ratios, its performance is quite limited for arbitrary bandwidth ratios. We propose strategies based on concatenation of layered source coding and channel coding with various cooperation protocols to improve the distortion exponent.

There is a close relation between the DMT performance of the underlying communication system and the achievable distortion exponent. Our techniques attempt to utilize the DMT curve most efficiently to maximize the achieved distortion exponent. In the LS strategy, where compressed source layers are transmitted progressively in time with the help of a cooperating relay, time and rate allocation among the layers is optimized. We show that this optimization is geometrically equivalent to climbing up the underlying DMT curve. An extension of LS is HLS, where an additional error message is transmitted in analog fashion without coding, directly from the source terminal. HLS strictly improves the distortion exponent compared to LS while the improvement decays with increasing bandwidth ratio. Both LS and HLS can easily be adapted to use any cooperative protocol for each layer. In BS multiple source layers are transmitted simultaneously by an appropriate power and rate allocation. Unlike LS and HLS, since BS superimposes layers, benefits of using a cooperation strategy with a better DMT is not immediately apparent, and one needs to consider the effect of successive decoding on the DMT. Coupling BS with a combination of direct transmission and cooperative relaying, we prove that it achieves the optimal distortion exponent of the half-duplex system for $b \leq 1$ and $b \geq 4$. We also analyze the distortion exponent of full-duplex relay system, and show that BS with full-duplex decode-and-forward protocol can achieve the optimal distortion exponent for all bandwidth ratios. Finally, we consider the availability of correlated side information at the destination and/or the relay terminals, and show that the side information does not effect the distortion exponents of the cooperation strategies we study.

Overall, our results show that layered source coding is an effective technique to improve the distortion exponent for communication over fading wireless channels, and to achieve a higher distortion exponent, we need to use better channel codes and cooperation strategies that improve DMT, and carefully match the code rates and cooperation scheme to the source coder.

Appendix I
Proof of Theorem 4.3

In [1], we proved that to solve (15)-(17) in the limit of infinite layers, equally dividing the total time among n layers, i.e., $t_i = 1/n$, is asymptotically optimal. Similarly, we can show that imposing $t_1 r_1 = \cdots = t_n r_n$ is asymptotically optimal as well.

Now assume that we have n layers that climb up to $\Delta = M$ on the tradeoff curve characterized by $d^*(r) = (1-r)/r$, where we impose the above equality constraint, i.e., $t_1 r_1 = \cdots = t_n r_n = M/bn$. Since $t_n r_n = M/bn = d^*(r_n)/b = (1-r_n)/(br_n)$, we have $bt_n = (M/n)(1+M/n)$. Continuing similarly, we get

$$\begin{aligned} bt_{n-1} &= \frac{M}{n}\left(1+\frac{2}{n}M\right), \\ &\vdots \\ bt_i &= \frac{M}{n}\left(1+\frac{n-i+1}{n}M\right), \\ &\vdots \\ bt_1 &= \frac{M}{n}\left(1+M\right). \end{aligned}$$

Summing up both sides, we get

$$\begin{aligned} \sum_{i=1}^{n} bt_i &= \frac{M}{n}\left(n+\frac{M}{n}\cdot\frac{n(n+1)}{2}\right) \\ b &= M+\frac{M^2(n+1)}{2n}. \end{aligned}$$

We solve this equation for M.

$$\begin{aligned} M &= -\frac{n}{n+1}+\frac{\sqrt{n^2+2bn(n+1)}}{n+1}, \\ &\xrightarrow{n\to\infty} -1+\sqrt{1+2b}. \end{aligned}$$

However, note that since the part of the tradeoff curve for $0 \leq r \leq 1/2$ is a straight line, above is valid only if $M \leq 1$, which corresponds to $b \leq 3/2$. For $b > 3/2$, we use the rest of the bandwidth ratio to climb up the straight line characterized by $d^*(r) = 2(1-r)$. This can be done using the results of [1] by assuming a bandwidth ratio of $\left(b-\frac{3}{2}\right)$ and adding 1 to the corresponding distortion exponent.

Appendix II
Proof of Theorem 4.4

The expected distortion for HLS with n layer source coding can be written as

$$\begin{aligned} \mathrm{ED}(\mathbf{R},\mathrm{SNR}) = & \textstyle\sum_{i=0}^{n-1} D_i^{HLS}\cdot\left(P_{out}^{i+1}-P_{out}^{i}\right) \\ & + \textstyle\int_{\mathcal{A}^c} D_a^{HLS}(h_1)p(\mathbf{h})d\mathbf{h}, \end{aligned} \tag{53}$$

where $P_{out}^0 = 0$, $P_{out}^{n+1} = 1$, $D_0^{HLS} = 0$,

$$D_i^{HLS} = D\left(\left(b-1\right)\sum_{k=1}^{i} t_k R_k\right) \text{ for } i = 1,\ldots,n \tag{54}$$

$$D_a^{HLS}(h_1) = \frac{D_n^{HLS}}{1+|h_1|^2\mathrm{SNR}}, \tag{55}$$

$\mathbf{h} = (h_1, h_2, h_3)$ is the channel state vector with probability density function $p(\mathbf{h})$, and $\mathcal{A}$ denotes the outage event for the n-th layer.

Suppose that the transmission rate of the n-th layer for the above strategy is $R = r\log\mathrm{SNR}$, where $r \leq 1$ is the multiplexing gain and $\mathcal{A}$ denotes the outage event for this layer. We also assume that the average received signal-to-noise ratio for the n-th coded layer and the analog transmitted portion are both SNR. Then we have

$$\int_{\mathcal{A}^c} D_a^{HLS}(h_1)p(\mathbf{h})d\mathbf{h} \quad \dot{\leq} \quad D_n^{HLS}\mathrm{SNR}^{-1}. \tag{56}$$

To prove this, we use Laplace's method following similar steps to the proof of Theorem 4 in [31]. Let $|h_1|^2 \doteq \mathrm{SNR}^{-\gamma}$, $|h_2|^2 \doteq \mathrm{SNR}^{-\beta}$, and $|h_3|^2 \doteq \mathrm{SNR}^{-\theta}$, and the joint probability density of γ, β, and θ is $p(\gamma,\beta,\theta)$. Then we can write

$$\begin{aligned} \int_{\mathcal{A}^c} D_a^{HLS}(h_1)p(\mathbf{h})d\mathbf{h} &\doteq D_n^{HLS}\int_{\mathcal{A}^c}\mathrm{SNR}^{-(1-\gamma)^+} \\ &\quad \cdot p(\gamma,\beta,\theta)d\gamma d\beta d\theta, \\ &\doteq D_n^{HLS}\int_{\mathcal{A}^c\cap\mathbb{R}^+}\mathrm{SNR}^{-(1-\gamma)^+} \\ &\quad \mathrm{SNR}^{-\gamma}\mathrm{SNR}^{-\beta}\mathrm{SNR}^{-\theta}d\gamma d\beta d\theta, \\ &\doteq D_n^{HLS}\mathrm{SNR}^{-\mu}, \end{aligned}$$

where

$$\begin{aligned} \mu &= \inf_{(\gamma,\beta,\theta)\in\mathcal{A}^c\cap\mathbb{R}^{3+}} (1-\gamma)^+ + \gamma + \beta + \theta, \\ &\geq 1, \end{aligned} \tag{57}$$

since $(1-\gamma)^+ + \gamma \geq 1$ for any $\gamma \geq 0$.

Eqn. (56) tells us that the SNR exponent of the last term is at least 1. We will assume that it behaves as SNR^{-1} to find a lower bound to the overall distortion exponent. Using $d^*(r)$, the DMT curve of the particular cooperation protocol utilized for the source layers, and defining $d^*(r_{k+1}) = 1$, we can write the high SNR approximation for the expected distortion of HLS as

$$\begin{aligned}
\mathrm{ED}(\mathbf{R}, \mathbf{t}, \mathrm{SNR}) &\doteq \sum_{i=0}^{n-1} \left[\mathrm{SNR}^{-d^*(r_{i+1})} - \mathrm{SNR}^{-d^*(r_i)}\right] \\
&\cdot \mathrm{SNR}^{-(b-1)\sum_{k=1}^{i} t_k r_k} + \mathrm{SNR}^{-(b-1)\sum_{k=1}^{n} t_k r_k - 1} \\
&\doteq \sum_{i=0}^{n} \mathrm{SNR}^{-d^*(r_{i+1}) - (b-1)\sum_{k=1}^{i} t_k r_k} \\
&\doteq \mathrm{SNR}^{\max_{0\le i\le n}\left[-d^*(r_{i+1}) - (b-1)\sum_{k=1}^{i} t_k r_k\right]}.
\end{aligned}$$

An analysis similar to Section IV-A with infinite number of layers gives us the distortion exponents for AF/DF and DDF protocols.

Appendix III
Successive Decoding Diversity Gain of AF

We first find the maximum mutual information of the j-th layer for the AF protocol assuming all previous layers have been decoded and subtracted. Layers $j+1, \ldots, n$ will be treated as noise. First, we define

$$\overline{x}_{s,j} \triangleq \sum_{i=j}^{n} \sqrt{\mathrm{SNR}_j}\, x_{s,j}.$$

Then the equivalent channel can be written as

$$\underbrace{\begin{bmatrix} y_d[i] \\ y_d[i+N/2] \end{bmatrix}}_{\mathbf{y}_d[i]} = \underbrace{\begin{bmatrix} h_1 \\ h_3\beta h_2 \end{bmatrix}}_{\mathbf{A}} \sqrt{\mathrm{SNR}_j}\, x_{s,j}[i] + \underbrace{\begin{bmatrix} h_1 & 0 & 1 & 0 \\ h_3\beta h_2 & h_3\beta & 0 & 1 \end{bmatrix}}_{\mathbf{B}} \underbrace{\begin{bmatrix} \overline{x}_{s,j+1}[i] \\ z_r[i] \\ z_d[i] \\ z_d[i+N/2] \end{bmatrix}}_{\mathbf{z}[i]}, \quad (58)$$

where the noise covariance matrix is $E[\mathbf{z}\mathbf{z}^\dagger] = diag(\overline{\mathrm{SNR}}_{j+1}, 1, 1, 1)$. The mutual information for source layer j is maximized for circularly symmetric complex Gaussian input, and is found to be

$$I_{AF,j} = \log\det(\mathbf{I} + \mathrm{SNR}_j \mathbf{A}\mathbf{A}^\dagger (\mathbf{B}E[\mathbf{z}\mathbf{z}^\dagger]\mathbf{B}^\dagger)^{-1}).$$

We can find

$$\begin{aligned}
\mathbf{A}\mathbf{A}^\dagger &= \begin{bmatrix} |h_1|^2 & \beta h_1 h_2^* h_3^* \\ \beta h_1^* h_2 h_3 & \beta^2 |h_2|^2 |h_3|^2 \end{bmatrix}, \\
\mathbf{B}\mathrm{E}[\mathbf{z}\mathbf{z}^\dagger]\mathbf{B}^\dagger &= \\
&\begin{bmatrix} |h_1|^2 \overline{\mathrm{SNR}}_{j+1} + 1 & \beta h_1 h_2^* h_3^* \overline{\mathrm{SNR}}_{j+1} \\ \beta h_1^* h_2 h_3 \overline{\mathrm{SNR}}_{j+1} & \beta^2 |h_3|^2 (|h_2|^2 \overline{\mathrm{SNR}}_{j+1} + 1) + 1 \end{bmatrix}.
\end{aligned}$$

We define

$$\begin{aligned}
w &\triangleq |h_1|^2 + |h_1|^2|h_2|^2\mathrm{SNR} + |h_3|^2(|h_1|^2 + |h_2|^2)\mathrm{SNR}, \\
u &\triangleq 1 + |h_2|^2\mathrm{SNR} + |h_3|^2\mathrm{SNR}.
\end{aligned}$$

Then the successive decoding outage probability for layer j can be written as

$$\begin{aligned}
P_{out}^{d,j} &= Pr\left\{\frac{1}{2} I_{AF,j} < R_j\right\}, \\
&= Pr\left\{\frac{w\overline{\mathrm{SNR}}_j + u}{w\overline{\mathrm{SNR}}_{j+1} + u} < \mathrm{SNR}^{2r_j}\right\}, \\
&= Pr\{w(\mathrm{SNR}^{1-2r_1-\cdots-2r_{j-1}-\epsilon_{j-1}} - \\
&\quad \mathrm{SNR}^{1-2r_1-\cdots-2r_{j-1}-2r_j-\epsilon_j}) < u(\mathrm{SNR}^{2r_j} - 1)\}, \\
&\doteq Pr\left\{w\mathrm{SNR}^{1-2r_1-\cdots-2r_{j-1}-\epsilon_j} < u\mathrm{SNR}^{2r_j}\right\}, \\
&\doteq Pr\left\{w\mathrm{SNR}/u < \mathrm{SNR}^{2(r_1+\cdots+r_j+\epsilon_j)}\right\}, \\
&\doteq Pr\left\{|h_1|^2\mathrm{SNR} + \frac{|h_2|^2|h_3|^2\mathrm{SNR}^2}{1 + |h_2|^2\mathrm{SNR} + |h_3|^2\mathrm{SNR}} \right. \\
&\quad \left. < \mathrm{SNR}^{2(r_1+\cdots+r_j)}\right\},
\end{aligned}$$

as $\epsilon_j \to 0$. Comparing this last equality with the outage probability expression of the AF protocol for a single stream in [5], we conclude that in the high SNR regime,

$$P_{out}^{d,j} \doteq \mathrm{SNR}^{-d_{AF}^*(r_1+\cdots+r_j)}, \quad (59)$$

and we have

$$\begin{aligned}
\bar{P}_{out}^{d,j} &\doteq \max(\bar{P}_{out}^{d,j-1}, P_{out}^{d,j}), \quad (60)\\
&\doteq P_{out}^{d,j}. \quad (61)
\end{aligned}$$

Therefore $d_{sd}^{AF}(r_i) = d_{AF}^*(r_1 + \cdots + r_i)$, and the diversity-multiplexing tradeoff curve of the AF protocol with the proposed exponential power allocation is successively refinable.

Appendix IV
Proof of Lemma 4.6

The proof closely resembles the proofs in [6], so we will briefly give the outline. Outage probability at the destination is the sum of two probabilities: probability of outage given the relay can decode the message at the end of the first portion and probability of outage given the relay cannot decode:

$$\begin{aligned}
P_{out} &= Pr(\text{outage} \mid \text{relay decodes})Pr(\text{relay decodes}) \\
&\quad + Pr(\text{outage}|\text{relay cannot decode}) \\
&\quad \cdot Pr(\text{relay cannot decode}).
\end{aligned}$$

Then we have

$$\begin{aligned}
&Pr(\text{outage} \mid \text{relay decodes}) = Pr\{\alpha\log(1 + |h_1|^2\mathrm{SNR}) + \\
&\quad (1-\alpha)\log(1 + |h_3|^2\mathrm{SNR}) \le r\log\mathrm{SNR}\}, \\
&Pr(\text{outage} \mid \text{relay cannot decode}) = \\
&\quad Pr\{\alpha\log(1 + |h_1|^2\mathrm{SNR}) \le r\log\mathrm{SNR}\}.
\end{aligned}$$

We also have

$$\begin{aligned}
&Pr(\text{relay can not decode}) = \\
&\quad Pr\{\alpha\log(1 + |h_2|^2\mathrm{SNR}) \le r\log\mathrm{SNR}\}.
\end{aligned}$$

Let $|h_1|^2 = \mathrm{SNR}^{-\gamma}$, $|h_2|^2 = \mathrm{SNR}^{-\beta}$, and $|h_3|^2 = \mathrm{SNR}^{-\theta}$. Now we can show that

$$\begin{aligned}
&Pr(\text{outage}|\text{relay cannot decode})Pr(\text{relay cannot decode}) \\
&\doteq \mathrm{SNR}^{-d_1(r)},
\end{aligned}$$

where $d_1(r) = 2(1 - \frac{r}{\alpha})$ for $r < \alpha$ and 0 elsewhere. On the other hand,

$$Pr(\text{outage} \mid \text{relay decodes })Pr(\text{relay decodes}) \doteq \text{SNR}^{-d_2(r)}$$

where

$$d_2(r) = \inf_{(\gamma,\theta)\in\mathcal{A}} \gamma + \theta,$$

with

$$\mathcal{A} = \{(\gamma, \theta) \in \mathbb{R}^{2+} : \alpha(1-\gamma)^+ + (1-\alpha)(1-\theta)^+ \le r\}.$$

We can show that

$$d_2(r) = \begin{cases} 2 - \frac{r}{1-\alpha} & \text{if} \quad r \le 1-\alpha \\ \frac{1-r}{\alpha} & \text{if} \quad r > 1-\alpha. \end{cases} \tag{62}$$

Then we have $d^*_{gDF}(r) = \min(d_1(r), d_2(r))$.

Appendix V
Proof of Lemma 4.7

Suppose the first $k-1$ layers have been decoded at the relay. Let $\bar{x}_{s,k} = \sum_{i=k}^n \sqrt{\text{SNR}_i^s} x_{s,i}$ and $\bar{y}_{r,k} = h_2\bar{x}_{s,k} + z_r$. We find the mutual information as

$$\begin{aligned} \mathcal{I}_{s\to r} &\triangleq \alpha\mathcal{I}(\sqrt{\text{SNR}_k^s}x_{s,k}; \bar{y}_{r,k}), \\ &= \alpha\mathcal{I}(\sqrt{\text{SNR}_k^s}x_{s,k}, \bar{x}_{s,(k+1)}; \bar{y}_{r,k}) \\ &\quad -\alpha\mathcal{I}(\bar{x}_{s,k+1}; \bar{y}_{r,k} | \sqrt{\text{SNR}_k^s}x_{s,k}), \\ &= \alpha\log(1 + \overline{\text{SNR}}_k^s \|h_2\|^2) \\ &\quad -\alpha\log(1 + \overline{\text{SNR}}_{k+1}^s \|h_2\|^2), \\ &= \alpha\log\left(\frac{1 + \overline{\text{SNR}}_k^s\|h_2\|^2}{1 + \overline{\text{SNR}}_{k+1}^s\|h_2\|^2}\right). \end{aligned}$$

Similar to the successive decoding outage events at the destination defined in Section IV-C, we define the following successive decoding outage events for layer k at the relay terminal,

$$\begin{aligned} \mathcal{O}_k^r &= \{h_2 : \mathcal{I}_{s\to r} < R_k\}, \\ \bar{\mathcal{O}}_k^r &= \bigcup_{i=1}^k \mathcal{O}_k^r, \end{aligned}$$

and the corresponding outage probabilities

$$\begin{aligned} P_{out}^{r,k} &= Pr\{h_2 : h_2 \in \mathcal{O}_k^r\}, \\ \bar{P}_{out}^{r,k} &= Pr\{h_2 : h_2 \in \bar{\mathcal{O}}_k^r\}. \end{aligned}$$

Now we look at the outage event at the destination for layer k assuming that the previous layers have been successfully decoded. Let $\bar{x}_{r,k} = \sum_{i=k}^n \sqrt{\text{SNR}_i} x_{r,i}$ and $\bar{y}_{d,k} = h_1\bar{x}_{s,k} + h_3\bar{x}_{r,k} + z_r$. The outage event at the destination will depend on the outage event at the relay for layer k. Let τ_k be the random variable denoting the outage event at the relay, i.e., we have $\tau_k = 1$ if layer k can be decoded at the relay, and $\tau_k = 0$ otherwise. Then we can write

$$\begin{aligned} \mathcal{I}_{s,r\to d} &\triangleq \alpha\mathcal{I}(\sqrt{\text{SNR}_k^s}x_{s,k}; \bar{y}_{d,k}) \\ &\quad +(1-\alpha)s_k\mathcal{I}(\sqrt{\text{SNR}_k^r}x_{r,k}; \bar{y}_{d,k}), \end{aligned}$$

where

$$\mathcal{I}(\sqrt{\text{SNR}_k^r}x_{r,k}; \bar{y}_{d,k}) = \log\left(\frac{1 + \overline{\text{SNR}}_k^r\|h_3\|^2}{1 + \overline{\text{SNR}}_{k+1}^r\|h_3\|^2}\right).$$

Note here that, it is possible that the relay decodes layer k but not all the n layers, and hence not transmit anything for those uncoded layers. However, in the mutual information expression we assume all the following layers act as noise, which in turn gives a lower bound to the mutual information. Then the successive decoding outage event for layer k at the destination terminal can be written as,

$$\mathcal{O}_k^d = \{(h_1, h_2, h_3) : \mathcal{I}_{s,r\to d} < R_k\}.$$

Recall that $P_{out}^{d,k}$ denotes the probability of outage for layer k at the destination, assuming that the decoder successfully decodes and subtracts the previous $k-1$ layers. On the other hand, $\bar{P}_{out}^{d,k}$ is the overall outage probability of layer k for successive decoding.

For $k = 1, \ldots, n$ we have

$$\begin{aligned} P_{out}^{d,k} &= Pr\Big\{\alpha\mathcal{I}(\sqrt{\text{SNR}_k^s}x_{s,k}; \bar{y}_{d,k}) \\ &\quad +(1-\alpha)\mathcal{I}(\sqrt{\text{SNR}_k^r}x_{r,k}; \bar{y}_{d,k}) < R_k\Big\}(1 - \bar{P}_{out}^{r,k}) \\ &\quad +Pr\Big\{\alpha\mathcal{I}(\sqrt{\text{SNR}_k^s}x_{s,k}; \bar{y}_{d,k}) < R_k\Big\}\bar{P}_{out}^{r,k} \end{aligned}$$

Let $|h_1|^2 = \text{SNR}^{-\gamma}$, $|h_2|^2 = \text{SNR}^{-\beta}$, and $|h_3|^2 = \text{SNR}^{-\theta}$. Using the source and the relay power allocation, for $k = 1, \ldots, n-1$, we have

$$\begin{aligned} \mathcal{O}_k^r = &\Big\{\beta : \alpha\Big(\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \beta\big]^+ \\ &-\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \beta\big]^+\Big) < r_k\Big\}, \end{aligned}$$

and

$$\mathcal{O}_n^r = \Big\{\beta : \alpha\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{n-1}) - \epsilon_{n-1} - \beta\big]^+ < r_n\Big\}.$$

Similarly, we have, for $k = 1, \ldots, n-1$,

$$\begin{aligned} \mathcal{O}_k^d = &\Big\{(\gamma, \beta, \theta) : \alpha\Big(\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \gamma\big]^+ \\ &-\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \gamma\big]^+\Big) \\ &+(1-\alpha)\Big(\big[1 - \frac{1}{1-\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \theta\big]^+ \\ &-\big[1 - \frac{1}{1-\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \theta\big]^+\Big) < r_k \big| \tau_k = 0\Big\} \\ &\bigcup\Big\{(\gamma, \beta, \theta) : \alpha\Big(\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \gamma\big]^+ \\ &-\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \gamma\big]^+\Big) < r_k \big| \tau_k = 1\Big\}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{O}_n^d = &\{(\gamma, \beta, \theta) : \alpha\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{n-1}) - \epsilon_{n-1} - \gamma\big]^+ \\ &+ (1-\alpha)\big[1 - \frac{r_1 - \cdots - r_{n-1}}{1-\alpha} - \epsilon_{n-1} - \theta\big]^+ < r_n \big| \tau_k = 0\} \\ &\bigcup\{(\gamma, \beta, \theta) : \\ &\alpha\big[1 - \frac{r_1 - \cdots - r_{n-1}}{\alpha} - \epsilon_{n-1} - \gamma\big]^+ < r_n \big| \tau_k = 1\}, \end{aligned}$$

Using Laplace's method, for the above multiplexing gain and power allocation rule among the layers, high SNR behavior of the outage probabilities $P_{out}^{r,k}$ for $k = 1, \dots, n$ are

$$P_{out}^{r,k} \doteq \text{SNR}^{-d_{sd}^{s\to r}(r_k)},$$

where

$$d_{sd}^{s\to r}(r_k) = \inf_{\beta \in \mathcal{O}_k^r \cap \mathbb{R}^+} \beta,$$

which leads to, as $\epsilon_{n-1} \to 0$,

$$d_{sd}^{s\to r}(r_k) = \left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_k)\right)^+.$$

Similarly, the successive decoding outage probability at the destination given the relay is in outage has the same high SNR outage behavior

$$Pr(\text{outage}|\text{relay cannot decode layer } k) \doteq \text{SNR}^{-d_{sd}^{s\to r}(r_k)}.$$

For the successive decoding outage probability at the destination given the relay decodes the message, we have

$$Pr(\text{outage} \mid \text{relay decodes layer } k) \doteq \text{SNR}^{-d_{sd}^{s,r\to d}(r_k)},$$

where

$$d_{sd}^{s,r\to d}(r_k) = \inf_{(\gamma,\beta,\theta)\in \mathcal{A}_k \cap \mathbb{R}^{3+}} \gamma + \theta,$$

with the set $\mathcal{A}_k$ defined as

$$\begin{aligned} \mathcal{A}_k \triangleq & \Bigg\{(\gamma,\beta,\theta) : \alpha\Bigg(\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \gamma\big]^+ \\ & -\big[1 - \frac{1}{\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \gamma\big]^+\Bigg) + \\ & (1-\alpha)\Bigg(\big[1 - \frac{1}{1-\alpha}(r_1 - \cdots - r_{k-1}) - \epsilon_{k-1} - \theta\big]^+ \\ & -\big[1 - \frac{1}{1-\alpha}(r_1 - \cdots - r_k) - \epsilon_k - \theta\big]^+\Bigg) < r_k | \gamma \le \theta\Bigg\}. \end{aligned}$$

Then as $\epsilon_{k-1} \to 0$, $d_{sd}^{s,r\to d}(r_k)$ can be found as

$$\begin{aligned} d_{sd}^{s,r\to d}(r_k) &= \left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_k)\right)^+ \\ &\quad + \left(1 - \frac{1}{1-\alpha}(r_1 + \cdots + r_{k-1})\right)^+ \end{aligned} \tag{63}$$

Overall, we have

$$P_{out}^{d,k} = \text{SNR}^{-d_{sd}(r_k)}, \tag{64}$$

and

$$\begin{aligned} \bar{P}_{out}^{d,k} &\doteq \max(\bar{P}_{out}^{d,k-1}, P_{out}^{d,k}), \\ &\doteq P_{out}^{d_{sd}(r_k)}. \end{aligned} \tag{65}$$

where

$$\begin{aligned} d_{sd}(r_k) &= \min\left\{2d_{sd}^{s\to r}(r_k), d_{sd}^{s,r\to d}(r_k)\right\}, \\ &= \min\Bigg\{2\left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_k)\right)^+, \\ & \left(1 - \frac{1}{\alpha}(r_1 + \cdots + r_k)\right)^+ \\ & + \left(1 - \frac{1}{1-\alpha}(r_1 + \cdots + r_{k-1})\right)^+\Bigg\}. \end{aligned} \tag{66}$$

Appendix VI
Proof of Theorem 4.10

We assign

$$\overline{\text{SNR}}_j = \text{SNR}^{1-2r_1-\cdots-2r_{j-1}}$$

for $j = 2, \dots, k+1$ and

$$\overline{\text{SNR}}_j = \text{SNR}^{1-2r_1-\cdots-2r_k - r_{k+1} - \cdots - r_{j-1}}$$

for $j = k+2, \dots, n$, where $r_i \in [0, 1/2]$ for $i = 1, \dots, k$ and $r_i \in [0,1]$ for $i = k+1, \dots, n$. Here the multiplexing gain vector $\mathbf{r}$ and $\overline{\text{SNR}}_j$ are defined as in Lemma 4.7. For this power allocation, we can prove that $d_{sd}^{DF-DT}(r_k) = d_{DF-DT}^*(r_1 + \cdots + r_k)$, that is, the diversity-multiplexing gain tradeoff of DT combined with DF is successively refinable. Details of this proof are given in Appendix VII.

Using the successive refinability of the DMT curve and Eqn. (31) we can solve for the optimal multiplexing gains. Since the first k layers are sent with cooperation using DF, we have $d_{sd}^{DF}(r_i) = d_{DF}^*(r_1 + \cdots + r_i) = 2 - 4(r_1 + \cdots + r_i)$ for $i = 1, \dots, k$, whereas for the remaining $(n-k)$ layers we have $d_{sd}^{DT}(r_i) = d_{DT}^*(r_1 + \cdots + r_i) = 1 - (r_1 + \cdots + r_i)$ for $i = k+1, \dots, n$. Here k, which describes the layer allocation among cooperation and direct transmission will be optimized as well.

In order to have all the terms in Eqn. (31) equal, we need

$$\begin{aligned} r_1 &= \frac{\left(\frac{4}{b} - 1\right)(2 - b - b^{n-k+1})}{(b-1)\left(\frac{b}{4}\right)^{k-1} b^{n-k+1} + \frac{4}{b}(4 - b - 3b^{n-k+1})}, \\ r_i &= \left(\frac{b}{4}\right)^{i-1} \text{ for } i = 1, \dots, k, \\ r_{k+1} &= \left[b + 3\frac{1 - \left(\frac{4}{b}\right)^k}{1 - \frac{4}{b}}\right] r_k - 1, \\ r_i &= b^{i-k-1} r_{k+1} \text{ for } i = k+2, \dots, n, \end{aligned}$$

For $k = \beta n$ with $0 < \beta < 1$, using (67) we get (46). Note that, as we let $n, k \to \infty$ we can approach any β as closely as desired.

Now, if we have $0 < b < 1$, then we also have $b < 4^\beta$ for any $\beta \ge 0$. In the limit of infinite layers, i.e., as $n \to \infty$, we have $\Delta = b$. We already know that, for this range of bandwidth ratios, $\beta = 0$ (direct transmission) would also give us the optimal distortion exponent of $\Delta = b$ as $n \to \infty$. For $1 \le b < 4$, fix β such that $4^\beta > b$ and $\beta < 1$. In this case we have $\Delta = (b+2)/3$ as $n \to \infty$. Note that, this corresponds to having both cooperative and direct layers. Finally, for $b \ge 4$, we have $b/4^\beta > b^{(1-\beta)}$, and when we take the limit as $n \to \infty$, we get $\Delta = 2$. Note that this would hold for any $\beta > 0$, while $\beta = 1$ (DF cooperation) would already be optimal in the limit of infinite layers.

Appendix VII
Successive Decoding Diversity Gain of DF-DT

First consider layers $j = k+1, \dots, n$. If the destination can decode all the previous layers, since the power allocation is the same as DT for which the DMT is successively refinable, we find $d_{sd}^{DF-DT}(r_j) = 1 - (r_1 + \cdots + r_j)$.

For layers $j = 1, \ldots, k$, when there is outage at the relay, the outage probability expression at the destination is same as transmitting all layers with DF protocol, since the latter layers have the same total power and all are transmitted only through the direct channel. However, when layer j, for $j = 1, \ldots, k$, can be decoded at the relay, the destination will combine both the source and the relay signals for decoding. Assume that the relay transmits all the layers after j up to k as well. This will only degrade the performance for layer j as those layers act as noise when decoding that layer. Now, when decoding layer j at the destination, the received signal at the destination in the second half will be composed of layers $j, \ldots, k$ transmitted from the relay, and layers $k+1, \ldots, n$ transmitted from the source directly.We have

$$\begin{aligned} y_d[i] &= h_3 \sum_{m=j}^{k} \sqrt{\mathrm{SNR}_m^r} x_{r,m}[i] \\ &\quad + h_1 \sum_{m=j+1}^{n} \sqrt{\mathrm{SNR}_m^r} x_{s,m}[i] + z_d[i], \end{aligned}$$

for $j = N/2+1, \ldots, N$. Then for the second half of the time slot, the mutual information between the relay codeword for layer j and the received signal at the destination is found as

$$\begin{aligned} &\log\left(1 + \frac{|h_3|^2 \mathrm{SNR}_j}{1 + |h_3|^2 \sum_{m=j+1}^{k} \mathrm{SNR}_m + |h_1|^2 \overline{\mathrm{SNR}}_{k+1}}\right) \\ &= \log\left(\frac{1 + |h_3|^2 \sum_{m=j}^{k} \mathrm{SNR}_m + |h_1|^2 \overline{\mathrm{SNR}}_{k+1}}{1 + |h_3|^2 \sum_{m=j+1}^{k} \mathrm{SNR}_m + |h_1|^2 \overline{\mathrm{SNR}}_{k+1}}\right), \\ &= \log\left(\frac{1 + |h_3|^2 \overline{\mathrm{SNR}}_j + (|h_1|^2 - |h_3|^2) \overline{\mathrm{SNR}}_{k+1}}{1 + |h_3|^2 \overline{\mathrm{SNR}}_{j+1} + (|h_1|^2 - |h_3|^2) \overline{\mathrm{SNR}}_{k+1}}\right), \\ &\geq \log\left(\frac{1 + |h_3|^2 \overline{\mathrm{SNR}}_j}{1 + |h_3|^2 \overline{\mathrm{SNR}}_{j+1}}\right), \end{aligned}$$

if $|h_1|^2 \leq |h_3|^2$, that is, for this case, the mutual information is at least as much as transmitting all layers from the relay terminal, as in the pure cooperation protocol. Hence, if the condition $|h_1|^2 \leq |h_3|^2$ is satisfied, then transmitting layers $k+1, \ldots, n$ from the source in the second half, instead of transmitting from the relay can only increase the mutual information for the layers $1, \ldots, k$, as the source-destination channel is weaker, and hence has lesser interference compared to the relay-destination channel. If we have $|h_1|^2 \geq |h_3|^2$, the mutual information term coming from the first half of the time slot would be dominant in the high SNR regime. In this case, since the second half would have no effect on outage event anyway, the high SNR behavior would be equivalent to transmitting all layers with cooperation.

We conclude that the successive decoding outage probability for the first k layers which are transmitted by simple cooperation is the same as when all the layers are transmitted by cooperation. Hence, the DMT curve is successively refinable for these layers as well, and we have $d_{sd}^{DF}(r_j) = 2 - 4(r_1 + \cdots + r_j)$ for $j = 1, \ldots, k$.

Appendix VIII
Proof of Theorem 5.2

The source terminal uses block Markov superposition encoding and the destination does backward decoding [43], [44]. Total N uses of the channel, during which the fading is constant, is divided into $B+1$ blocks, each of length $N/(B+1)$ channel uses. Let $w_{i,b}$ be the message corresponding to layer i in block b, which is compressed at rate bR_i with $i = 1, \ldots, n$, $b = 1, \ldots, B$.

In the first block, the source transmits a superposition of the codewords corresponding to all layers for $b = 1$ given by the scaled versions of $x_{s,1}(1, w_{1,1}), \ldots, x_{s,n}(1, w_{n,1})$, while the relay transmits a known signal. The relay then tries to decode as many layers as possible. The number of layers that the relay can decode depends on the fading level of the source-relay channel and can be considered as constant for all the $B+1$ blocks, since the channel is constant. In the next block, the relay transmits the layers it decodes, superimposing them with the same power allocation as the source, that is, the relay transmits a superposition of scaled versions of $x_{r,1}(w'_{1,1}), \ldots, x_{r,k}(w'_{k,1})$. Here $w'_{j,1}$ denotes the estimate of the relay and is equal to $w_{j,1}$ for $j = 1, \ldots, k$ with high probability as these layers are not in outage. Both the source and the relay use Gaussian codebooks.

In the second and the following blocks, the source transmits a superposition of codewords for all layers using a doubly indexed codebook based on the current and the previous block messages (see Fig. 12). The destination combines the information from the source and the relay, and tries to decode as many layers as possible starting from block $B+1$ and going backwards. For the layers that the relay cannot decode, destination decodes them based on its received signal from the source terminal only. We assume B is large and $N/(B+1)$ is integer. Both are valid based on the infinite N assumption. Further details of encoding and decoding structure can be found in [44].

Let the received signals at the relay and the destination for block b be $y_{r,b}$ and $y_{d,b}$, respectively. We have

$$y_{r,b} = h_2 \sum_{i=1}^{n} \sqrt{\mathrm{SNR}_i} x_{s,i}(w_{i,b-1}, w_{i,b}) + z_r, \quad (67)$$

$$\begin{aligned} y_{d,b} &= h_1 \sum_{i=1}^{n} \sqrt{\mathrm{SNR}_i} x_{s,i}(w_{i,b-1}, w_{i,b}) \\ &\quad + h_3 \sum_{i=1}^{n} \sqrt{\mathrm{SNR}_i} x_{r,i}(w'_{i,b-1}) + z_d, \end{aligned} \quad (68)$$

where z_r and z_d are independent additive white Gaussian noise vectors with i.i.d. components of zero mean and variance $1/2$, $\sqrt{\mathrm{SNR}_i} x_{s,i}$, $\sqrt{\mathrm{SNR}_i} x_{r,i}$ are the source and the relay signals, respectively, corresponding to the i-th source layer. We assume $x_{r,i}$ is 0 if the layer i cannot be decoded at the relay. Due to power constraints at the source and the relay, we have $||x_{s,i}||^2 = ||x_{r,i}||^2 = 1$, while $\sum_{j=1}^{n} \mathrm{SNR}_j \leq \mathrm{SNR}$.

For each block, both the relay and the destination use successive decoding starting from the first layer. For the power allocation

$$\overline{\mathrm{SNR}}_k = \mathrm{SNR}^{1 - r_1 - \cdots - r_{k-1} - \epsilon_{k-1}}, \quad (69)$$

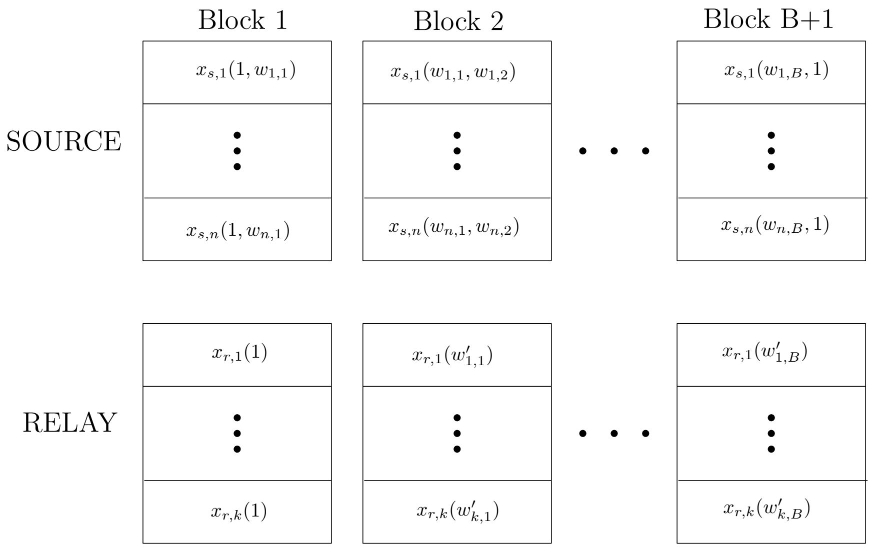

Fig. 12. Block Markov encoding for full duplex relay BS. In the figure, it assumed that the relay can decode, and thus can forward, up to layer k. Also $w'_{b,j}$ denotes the relay's estimate of layer j in block b.

for some $0 < \epsilon_1 < \cdots < \epsilon_{n-1}$ where $\overline{\text{SNR}}_k$ was defined in Section IV-C, we show in Appendix IX that $d_{sd}(r_k) = 2(1 - r_1 - \cdots - r_k)$. Therefore, the full-duplex DMT is successively refinable.

Assigning the multiplexing gains as below, all the terms in Eqn. (31) will be the same.

$$r_1 = \frac{1 - b/2}{1 - (b/2)^{n+1}}, \quad (70)$$

$$r_i = \frac{b}{2} r_{i-1}, \text{ for } i = 2, 3, \ldots, n, \quad (71)$$

and we find the distortion exponent given in (50).

APPENDIX IX
SUCCESSIVE DECODING DIVERSITY GAIN OF FULL-DUPLEX DF

The source codeword $x_{s,i}$ for layer i can be considered as the superposition of the two codewords, one for the current message $x^2_{s,i}$, and the other for the previous block's message $x^1_{s,i}$, where we dropped the dependence on the messages for simplicity. Both codewords are formed using Gaussian distribution. Without loss of generality, we assume $||x^1_{s,i}||^2 = ||x^2_{s,i}||^2 = 1/2$.

At each block, the relay only tries to decode the layers that it could decode in the previous block. Then, for each block, when decoding the message corresponding to a source layer in that block, the relay already knows the previous block's message corresponding to that layer. Thus the mutual information related to the new message at layer k is

$$\log\left(\frac{1 + \frac{1}{2}\overline{\text{SNR}}_k \|h_2\|^2}{1 + \overline{\text{SNR}}_{k+1}\|h_2\|^2}\right), \quad (72)$$

where we can omit the $1/2$ term in the numerator as that has no effect on the high SNR behavior.

The successive decoding outage event for layer k at the relay terminal is then found as,

$$\mathcal{O}^r_k = \{h_2 : \log\left(\frac{1 + \overline{\text{SNR}}_k \|h_2\|^2}{1 + \overline{\text{SNR}}_{k+1}\|h_2\|^2}\right) < R_k\}, \quad (73)$$

$$\bar{\mathcal{O}}^r_k = \bigcup_{i=1}^{k} \mathcal{O}^r_k, \quad (74)$$

where the corresponding outage probabilities are defined as in Appendix IV.

$$P^{r,k}_{out} = \{h_2 : h_2 \in \mathcal{O}^r_k\}, \quad (75)$$

$$\bar{P}^{r,k}_{out} = \{h_2 : h_2 \in \bar{\mathcal{O}}^r_k\}. \quad (76)$$

If the relay can decode up to layer k in the first block, then it can decode k layers in the following blocks as well, as it has a perfect knowledge of the previous block's information up to layer k and the mutual information corresponding to layers $1, \ldots, k$ are the same as in the first block.

The destination uses backward decoding. Hence it starts decoding from the last block and decodes as many layers as possible. Since in this block, the source terminal does not send any new information, the destination receives layers decoded by the relay from both terminals, and the remaining layers only from the source terminal. Then it moves backwards to the next block. If the destination can decode up to some layer j in the previous block, then knowing the message for these layers for the previous block, it can decode the same number of layers in the next block as well, since the mutual information for layers $1, \ldots, j$ in the next block would be equal to the previous block.

Let $\bar{x}_{r,j} = \sum_{i=j}^{n} \sqrt{\text{SNR}_i} x_{r,i}$, $\bar{x}_{s,j} = \sum_{i=j}^{n} \sqrt{\text{SNR}_i} x_{s,i}$ and $\bar{y}_{d,j} = h_1 \bar{x}_{s,j} + h_3 \bar{x}_{r,j} + z_r$. The destination tries to decode the layers that it could decode in the previous block. For each block, when decoding the message corresponding to a certain source layer coming from both the source and the relay, the destination already knows the new information sent by the source related to the decoded layers due to backward decoding. Then the mutual information related to the message of layer j, that the receiver can gather is

$$\begin{cases} \log\left(\frac{1+\overline{\mathrm{SNR}}_j(\|h_1\|^2+\|h_3\|^2)}{1+\overline{\mathrm{SNR}}_{j+1}(\|h_1\|^2+\|h_3\|^2)}\right) & \text{if relay decodes layer j,} \\ \log\left(\frac{1+\overline{\mathrm{SNR}}_j\|h_1\|^2}{1+\overline{\mathrm{SNR}}_{j+1}\|h_1\|^2}\right) & \text{otherwise.} \end{cases}$$

Note that, in the above mutual information expression, if layer k is decoded at the relay, when writing the noise term in the denominator inside the logarithm, we assume that all the following layers can also be decoded and forwarded by the relay, and hence act as noise at the destination when decoding layer j. This can only degrade the outage performance. Then outage event for layer j at the destination terminal can be written as,

$$\mathcal{O}_j^d = \{(h_1,h_2,h_3) : \mathcal{I}(\sqrt{\mathrm{SNR}_j}x_{s,j}^1, \sqrt{\mathrm{SNR}_j}x_{r,j}; \bar{y}_{d,j}) < R_j\},$$
$$\bar{\mathcal{O}}_j^d = \bigcup_{i=1}^{j} \mathcal{O}_j^d,$$

and the corresponding outage probabilities are the same as Eqn. (25)-(26).

For $j = 1,\ldots,n-1$ we have

$$P_{out}^{d,j} = Pr\left\{\left(\frac{1+\overline{\mathrm{SNR}}_j(\|h_1\|^2+\|h_3\|^2)}{1+\overline{\mathrm{SNR}}_{j+1}(\|h_1\|^2+\|h_3\|^2)}\right) < \mathrm{SNR}^{r_j}\right\} \cdot (1-\bar{P}_{out}^{r,j}) + Pr\left\{\left(\frac{1+\overline{\mathrm{SNR}}_j\|h_1\|^2}{1+\overline{\mathrm{SNR}}_{j+1}\|h_1\|^2}\right) < \mathrm{SNR}^{r_j}\right\}\bar{P}_{out}^{r,j}$$

and

$$P_{out}^{n} = Pr\left\{1+\overline{\mathrm{SNR}}_n(\|h_1\|^2+\|h_3\|^2) < \mathrm{SNR}^{r_n}\right\} \cdot (1-\bar{P}_{out}^{r,j}) + Pr\left\{1+\overline{\mathrm{SNR}}_n\|h_1\|^2 < \mathrm{SNR}^{r_n}\right\}\bar{P}_{out}^{r,j}.$$

Let $|h_1|^2 = \mathrm{SNR}^{-\gamma}$, $|h_2|^2 = \mathrm{SNR}^{-\beta}$, and $|h_3|^2 = \mathrm{SNR}^{-\theta}$. Then with the power allocation in (69), for $j = 1,\ldots,n-1$,

$$\begin{aligned} \mathcal{O}_j^r &= \{\beta : (1-r_1-\cdots-r_{j-1}-\epsilon_{j-1}-\beta)^+ \\ &\quad +(1-r_1-\cdots-r_j-\epsilon_j-\beta)^+ < r_j\}, \end{aligned}$$

and

$$\mathcal{O}_n^r = \{\beta : (1-r_1-\cdots-r_{n-1}-\epsilon_{n-1}-\beta)^+ < r_n\}.$$

Similarly, we have, for $j = 1,\ldots,n-1$,

$$\begin{aligned} \mathcal{O}_j^d = &\{(\gamma,\beta,\theta) : (1-r_1-\cdots-r_{j-1}-\epsilon_{j-1}-\min(\gamma,\theta))^+ \\ &+ (1-r_1-\cdots-r_j-\epsilon_j-\min(\gamma,\theta)))^+ < r_j | \beta \notin \bar{\mathcal{O}}_j^r\} \\ &\bigcup\{(\gamma,\beta,\theta) : (1-r_1-\cdots-r_{j-1}-\epsilon_{j-1}-\gamma)^+ \\ &+ (1-r_1-\cdots-r_j-\epsilon_j-\gamma)^+ < r_j | \beta \in \bar{\mathcal{O}}_k^r\}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{O}_n^d = &\{(\gamma,\beta,\theta) : \\ &(1-r_1-\cdots-r_{n-1}-\epsilon_{n-1}-\min(\gamma,\theta))^+ < r_n | \alpha_2 \notin \bar{\mathcal{O}}_n^r\} \\ &\bigcup\{(\gamma,\beta,\theta) : \\ &(1-r_1-\cdots-r_{n-1}-\epsilon_{n-1}-\gamma)^+ < r_n | \beta \in \bar{\mathcal{O}}_n^r\}, \end{aligned}$$

From Laplace's method [31], for the above multiplexing gain and power allocation among the layers, high SNR behavior of the outage probabilities $P_{out}^{d,j}$ for $j = 1,\ldots,n$.

$$P_{out}^{d,j} = \mathrm{SNR}^{-d_{sd}(r_j)}, \tag{77}$$

where

$$d_{sd}(r_j) = \inf_{(\gamma,\beta,\theta)\in\mathcal{O}_j^d\cap\mathbb{R}^{3+}} \gamma+\beta+\theta. \tag{78}$$

Now we have two minimizing (γ,β,θ) tuples, one for each set in $\mathcal{O}_j^d \cap \mathbb{R}^{3+}$ above. We have

$$\tilde{\gamma} = \tilde{\beta} = 1-r_1-\cdots-r_j-\epsilon_{j-1}, \text{ and } \tilde{\theta} = 0, \tag{79}$$

or,

$$\tilde{\gamma} = \tilde{\theta} = 1-r_1-\cdots-r_j-\epsilon_{j-1}, \text{ and } \tilde{\beta} = 0. \tag{80}$$

In both cases, the successive decoding diversity gain $d_{sd}(r_j)$ for $j = 1,\ldots,n$ as $\epsilon_j \to 0$ can be found as

$$d_{sd}(r_j) = 2(1-r_1-\cdots-r_j). \tag{81}$$

Due to successive decoding at the destination, we have

$$\begin{aligned} \bar{P}_{out}^{d,j} &\doteq \max(\bar{P}_{out}^{d,j-1}, P_{out}^{d,j}), & (82)\\ &\doteq P_{out}^{d,j}. & (83) \end{aligned}$$

We conclude that the successive decoding DMT for the exponential power allocation is equivalent to the DMT of a single data stream transmitted at multiplexing gain $(r_1+\cdots+r_k)$, that is, DMT of full-duplex DF is successively refinable.

References

[1] D. Gündüz, E. Erkip, "Joint source-channel codes for MIMO block fading channels," under revision, IEEE Transactions on Information Theory.

[2] J. N. Laneman, E. Martinian, G. W. Wornell, and J. G. Apostolopoulos, "Source-channel diversity for parallel channels," IEEE Transactions on Information Theory, vol. 51, no. 10, pp. 3518-3539, Oct. 2005.

[3] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, Part I and Part II," IEEE Trans. on Communications, vol. 51, pp. 1927-48, November 2003.

[4] A. Sendonaris, E. Erkip, and B. Aazhang, "User cooperation diversity, Part II: Implementation aspects and performance analysis," IEEE Trans. on Communications, vol. 51, pp. 1939-1948, November 2003.

[5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," IEEE Trans. Inform. Theory, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.

[6] K. Azarian, H. El Gamal, P. Schniter, "On the achievable diversity-multiplexing tradeoff in half-duplex cooperative channels," IEEE Trans. Info. Theory, vol. 51, no. 12, pp. 4152-4172, Dec. 2005.

[7] M. Yuksel, E. Erkip, "Multi-antenna cooperative wireless systems: A diversity multiplexing tradeoff perspective," to appear, IEEE Transactions on Information Theory.

[8] T. Holliday, A. Goldsmith, "Optimizing end-to-end distortion in MIMO systems," Proc. of IEEE International Symposium on Information Theory (ISIT), Adelaide, Australia, Sep. 2005.

[9] B. P. Dunn and J. N. Laneman, "Characterizing source-channel diversity approaches beyond the distortion exponent," in Proc. Allerton Conf. Communications, Control, and Computing, Monticello, IL, Sept. 2005.

[10] G. Caire, K. Narayanan , "On the SNR exponent of hybrid digital-analog space time coding," in Proceedings of Allerton Conference, Allerton, IL, September 2005.

[11] D. Gündüz, E. Erkip, "Joint source-channel cooperation: Diversity versus spectral efficiency," in Proc. of IEEE International Symposium on Information Theory (ISIT), Chicago, June 2004.

[12] D. Gündüz, E. Erkip, "Source and channel coding for cooperative relaying," in Proceedings of SPAWC, New York City, June 2005.

[13] D. Gündüz, E. Erkip, “Source and channel coding for quasi-static fading channels,” in Proceedings of 39th Asilomar Conf., Pacific Grove, California, November 2005.

[14] D. Gündüz, E. Erkip, “Distortion exponent of MIMO fading channels,” in Proceedings of IEEE Information Theory Workshop, Punta del Este, Uruguay, March 2006.

[15] D. Gündüz, E. Erkip, “Distortion exponent of parallel channels,” in Proc. of IEEE International Symposium on Information Theory (ISIT), Seattle, WA, July 2006.

[16] K. Bhattad, K. R. Narayanan, G. Caire, “On the distortion exponent of some layered transmission schemes”, in Proceedings of 40th Asilomar Conf., Pacific Grove, California, Nov. 2006.

[17] F. Etemadi, H. Jafarkhani, “Optimal layered transmission over quasi-static fading channels,” in Proc. of IEEE International Symposium on Information Theory (ISIT), July 2006.

[18] C. T. K. Ng, D. Gündüz, A. J. Goldsmith and E. Erkip, “Minimum expected distortion in Gaussian joint source-channel layered broadcast coding,” in Proc. of IEEE International Symposium on Information Theory (ISIT), Nice, France, June 2007.

[19] C. T. K. Ng, D. Gündüz, A. J. Goldsmith and E. Erkip, “Recursive power allocation in Gaussian joint source-channel layered broadcast coding,” in Proc. of IEEE International Conference on Communications (ICC 2007), Glasgow, Scotland, June 2007.

[20] X. Xu, D. Gündüz, E. Erkip, Y. Wang, “Layered cooperative source and channel coding,” IEEE International Conference on Communications (2):1200 - 1204, May. 2005.

[21] H. Y. Shutoy, D. Gündüz, E. Erkip and Y. Wang, “Cooperative source and channel coding for wireless multimedia communications,” to appear in IEEE Journal of Selected Topics In Signal Processing, special issue on Network Aware Multimedia Processing and Communications, 2007.

[22] A. Kwasinski, Z. Han, K. J. R. Liu, “Cooperative multimedia communications: Joint source coding and collaboration”, Proceeding of IEEE Global Telecommunications Conference (GLOBECOM), Saint Louis, MO, Nov.-Dec. 2005.

[23] A. Kwasinski, Z. Han, K. J. R. Liu, “Joint source coding diversity and cooperative diversity for multimedia communications”, Proc. of IEEE Inter. Workshop on Signal Processing Advances for Wireless Comm. (SPAWC), pp. 129-133, New York, NY, June 2005.

[24] A. Kwasinski, K. J. R. Liu, “Optimal unequal error protection with user cooperation for transmission of embedded source-coded images” ICIP 2006, Atlanta, GA, Oct. 2006.

[25] T. Cover, J. Thomas, *Elements of Information Theory,* Wiley and Sons, New York, 1991.

[26] P. Zador, “Topics in the asymptotic quantization of continuous random variables,” Bell Laboratories Technical Memorandum, 1966.

[27] N. Jayant and P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice Hall, 1984.

[28] T. Linder, R. Zamir, “On the asymptotic tightness of the Shannon lower bound”, IEEE Transactions on Information Theory, pp. 2026-2031, Nov. 1994.

[29] W. Equitz, T. Cover, “Successive refinement of information,” IEEE Transactions on Information Theory, 37(2): 269-275, Mar. 1991.

[30] L. Lastras, T. Berger, “All sources are nearly successively refinable,” IEEE Transactions on Information Theory, 47(3): 918-926, March 2000.

[31] L. Zheng, D. Tse, “Diversity and multiplexing: A fundamental tradeoff in multiple antenna channels,” IEEE Transactions on Information Theory, vol. 49, pp. 1073-96, May 2003.

[32] U. Mittal, N. Phamdo, “Hybrid digital-analog joint source-channel codes for broadcasting and robust communications,” IEEE Transactions on Information Theory, vol. 48, pp. 1082-1102, May 2002.

[33] H. Coward, T. A. Ramstad, “Quantizer optimization in hybrid digital analog transmission of analog source signals,” in Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, Istanbul, Turkey, Jun. 2000.

[34] M. Skoglund, N. Phamdo, F. Alajaji, “Design and performance of VQ-based hybrid digital-analog joint source–channel codes,” IEEE Trans. Inf. Theory, vol. 48, no. 3, pp. 708720, Mar. 2002.

[35] M. Skoglund, N. Phamdo, F. Alajaji, “Hybrid digital-analog source-channel coding for bandwidth compression/expansion,” IEEE Transactions on Information Theory, Vol. 52, No. 8, pp. 3757-3763, August 2006.

[36] I. Kozintsev, K. Ramchandran, “Multiresolution joint source-channel coding using embedded constellations for power-constrained time-varying channels,” In Proc. of IEEE Inter. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP), Atlanta, GA, May 1996.

[37] S. Shamai, “A broadcast strategy for the Gaussian slowly fading channel,” Proceedings of ISIT, p. 150, June 1997.

[38] T. M. Cover, “Broadcast channels,” IEEE Transactions on Information Theory, vol. IT-18 pp. 2-14, January 1972.

[39] S. Shamai, A. Steiner, “A broadcast approach for a single-user slowly fading MIMO channel,” IEEE Transactions on Information Theory, October 2003.

[40] S. Diggavi, D. Tse, “Fundamental limits of diversity-embedded codes over fading channels,” Proceedings ISIT, Adelaide, Australia, September 2005.

[41] S. Diggavi, D.N.C. Tse, “On opportunistic codes and broadcast codes with degraded message sets,” In Proc. of IEEE Information Theory Workshop, Punta del Este, Uruguay, March 2006.

[42] N. Prasad, M. K. Varanasi, ”High performance static and dynamic cooperative communication protocols for the half duplex fading relay channel,” in Proc. of IEEE Globecom Conference, San Fransisco, CA, Nov. 2006.

[43] TM Cover and A. El Gamal, “Capacity theorems for the relay channel,” IEEE Trans. Inform. Theory, vol. IT-25, no. 5, pp. 572-584, Sept. 1979.

[44] G. Kramer, M. Gastpar, P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” IEEE Transactions on Information Theory, 51(9):3037-3063, September 2005.

[45] S. Draper, G. Wornell, “Side information aware coding strategies for sensor networks,” IEEE J. Select. Areas Commun., vol. 22, Aug 2004.

[46] D. Gündüz, C. T. K. Ng, E. Erkip and A. J. Goldsmith “Source transmission over relay channel with correlated relay side information,” in Proc. of IEEE International Symposium on Information Theory (ISIT 2007), Nice, France, June 2007.

[47] D. Gündüz, E. Erkip, “Reliable cooperative source transmission with side information,” IEEE Information Theory Workshop (ITW 2007), Bergen, Norway, July 2007.

Deniz Gündüz [S'02] received the B.S. degree in Electrical and Electronics Engineering from Middle East Technical University, Ankara, Turkey, in 2002 and the M.S. degree in Electrical and Computer Engineering from Polytechnic University, Brooklyn, NY, in 2004. He is currently working toward his Ph.D. at Polytechnic University.

In 2004, he was a summer researcher in the laboratory of information theory (LTHI) at EPFL in Lausanne, Switzerland. His research interests lie in the areas of communications theory and information theory with special emphasis on joint source-channel coding, cooperative communications and cross-layer design.

Elza Erkip [S'93, M'96, SM'05] received the Ph.D. and M.S. degrees in Electrical Engineering from Stanford University, and the B.S. degree in Electrical and Electronics Engineering from Middle East Technical University, Turkey. She joined Polytechnic University in Spring 2000, where she is currently an Associate Professor of Electrical and Computer Engineering.

Dr. Erkip received the 2004 Communications Society Stephen O. Rice Paper Prize in the Field of Communications Theory and the NSF CAREER award in 2001. She is an Associate Editor of IEEE Transactions on Communications, a Publications Editor of IEEE Transactions on Information Theory and a Guest Editor of IEEE Signal Processing Magazine, Special Issue on Signal Processing for Multiterminal Communication Systems. She is the Technical Area Chair for the ”MIMO Communications and Signal Processing” track of 41st Annual Asilomar Conference on Signals, Systems, and Computers, and the Technical Program Co-Chair of 2006 Communication Theory Workshop. Her research interests are in wireless communications, information theory and communication theory.