Amplify-Forward and Decode-Forward: The Impact of Location and Capacity Contour

Meng Yu and Jing (Tiffany) Li

Department of Electrical and Computer Engineering
Lehigh University, Bethlehem, PA 18015.
Emails: {mey3,jingli}@ece.lehigh.edu

Hamid Sadjadpour

Baskin School of Engineering
University of California, Santa Cruz, CA 95064
Emails: hamid@soe.ucsc.edu

Abstract*— *Successful message relay, or the quality of the inter-user channel, is critical to fully realize the cooperative benefits promised by the theory. This in turn points out the importance of the relative location of the users. This paper investigates the impact of the location on the system capacity and outage probability for both amplify-forward (AF) and decode-forward (DF) schemes. Signal attenuation is modeled using power laws and capacity is evaluated using the max-flow min-cut theory. The resemblance and difference between cooperative systems and multi-input multi-output (MIMO) systems are also discussed. Finally a capacity contour for DF, the more popular mode of the two, is provided to facilitate the derivation of engineering rules.

I. INTRODUCTION

Aside from temporal and frequency diversities, spatial diversity is another technique to mitigate the deterioration caused by fading. Due to the limitation on the size of mobile terminals, multiple antennas are not always practical. As a remedy to this, user cooperation has been proposed [1][2], where multiple users share antennas to form a virtual antenna array and obtain spatial diversity.

Aiming at increasing the channel capacity and/or decreasing the outage probability, several interesting cooperative protocols have been proposed (e.g. [2][3]). Among them, amplify-forward (AF) and decode-forward (DF) are the two fundamental forwarding modes. Their qualities have been studied by many researchers both from the information theoretic aspects and the practical aspects (e.g. [2], [4]). [5] evaluated their performances in practical wireless scenarios in general, and the inter-user outage case in particular. (By inter-user outage, we mean that the relay is unable to extract a clean copy of the source data.) It was shown [5] that (1) the inter-user outage happens at a non-negligible probability even with decent channel code protection; for example, on block Rayleigh fading channels with an inter-user signal-to-noise ratio (SNR) of 10-22 dB, inter-user outage happens at a probability of 10.4%-1.06% even with the protection of a $(3000, 2000)$ random low-density parity-check (LDPC) code; (2) when inter-user outage happens, both AF and DF perform badly with an effective diversity order of only 1; and (3) the overall system performance is to a large extend limited by this worst-case scenario. These results revealed that a high-quality inter-user channel is one key to realize the great benefits that user cooperation may offer. This in turn points out the importance of the location of the relay. In other words, a wireless node should judiciously choose its relay partner. Assuming the network is dense enough, what would be the desired location for the relay? Intuitively, when the relay gets close to the source, inter-user outage tends to diminish and the cooperative system tends to resemble a 2-by-1 multiple-input single-output (MISO) system; but how close is close? Further, since 2-by-1 and 1-by-2 systems are capacity comparable, does this suggest that a relay close to the destination would also work in favor? More generally, is there a symmetry or duality property in the relay system?

The purpose of this paper is to answer the above questions and to understand the effect caused by the location of the relay. [6] analyzed the performance of relay networks based on achievable rate region. This paper focuses on capacity evaluation, including the ergodic channel capacity and the outage capacity. Capacity by definition establishes limits on the performance of practical communication systems. These limits provide system benchmarks and reveal how much improvement is theoretically possible. Several researchers have studied the information-theoretic aspects of the two-transmitter one-destination wireless cooperative system but only for a few samples of fixed channel qualities. In this work, we also study the capacity as a function of geometry. A similar study was conducted for the Gaussian channels in [9]. We consider both DF and AF modes for the single-relay cooperative system on Rayleigh fading

This material is based on research supported by the National Science Foundation under Grant No. CCF-0430634 and by the Commonwealth of Pennsylvania, Department of Community and Economic Development, through the Pennsylvania Infrastrusture Technology Alliance (PITA).

channels using power law air propagation models. The system limits are first analyzed using the max-flow min-cut theory for different relay locations in several topologies. A capacity contour for DF, the more practical and useful mode of the two, is subsequently computed. The capacity contour clearly provides motivation and guidelines for choosing good partners in practical situations. It is also worth mentioning that, our analysis reveals a symmetry property in the capacity for the AF mode, but not in the DF mode. The reason behind it, particularly in terms of why DF does not mimic MIMO, is provided.

In the succeeding sections, we will brief on the system model and cooperation mode first. Then we will evaluate and analyze the ergodic capacity and outage probability affected by the relay at different positions. The result shows that when the relay is near the median line between the source and the destination, system will achieve high capacity with some exception in AF mode.

II. User Cooperation

A. System Model

In this paper we consider single relay wireless system. Let "home channel", "inter-user channel" and "relay channel" denote the channels between the source and the destination, the source and the relay, and the relay and the destination, respectively. Since user cooperation is most useful when channels are varying very slowly (i.e. hard to obtain time diversity in a single user channel), the channels are modeled as Rayleigh block fading channel, where channels remain constant for the duration of one round of user cooperation (2 consecutive time slots). Between channels and cooperation rounds, they are independent. Let h_{SD}, h_{SR} and h_{RD} denote the respective path gains. The general form of a signal received over a specific channel at time t is given by,

$$y(t) = \sqrt{E_s}h(t)x(t) + n(t), \tag{1}$$

where E_s is the signal energy, $h(t)$ the path gain, and $n(t)$ the additive white Gaussian noise (AWGN). The path gain follows

$$h(t) = \alpha\phi(t), \tag{2}$$

where $\phi(t)$ is a Rayleigh distributed random variable and α is the pathloss. We assume the square of pathloss inversely proportional to the power of the distance, i.e., $\alpha^2 = l^{-\delta}$. Where l is the distance between the transmitter and the receiver, and δ, an integer between $[2, 4]$, is the pathloss exponent[1]. Without loss of generality, we use normalized value or distribution for $\phi(t)$ and $n(t)$, i.e., $\phi(t)$ takes a unit-variance Rayleigh distribution and $n(t)$ takes a zero-mean and unit-variance Gaussian distribution ($N_0 = 1$).

[1]To ease the evaluation, we consider isotropic signal propagation model within each topological setup. In reality, however, the propagation model also depends on the environment and the transmission distance.

Among the various possible strategies of user cooperation (e.g. [2][3]), we consider half-duplex system, the simplest type, where after the source transmits a package in the first time slot, the relay forwards the message in the second time slot. We assume the channel side information (i.e. fading coefficients) is known to the respective receivers and the power is equally allocated in the two time slots.

B. Fundamental Cooperative Modes

At the first time slot, the signal received at the destination is

$$y_{D,1} = \sqrt{E_S}h_{SD}x_1 + n_0.$$

where n_0 denotes the zero-mean complex AWGN.

1) AF Mode: we assume that the power of the signal retransmitted at the relay node is scaled uniformly with respect to all the bits in the package, such that the average (re-)transmission energy per signal equals E_S.

In the second time slot, the signal received at the destination is

$$y_{D,2}^{AF} = h_{RD}|h_{SR}|\sqrt{\frac{E_S^2}{E_S|h_{SR}|^2 + N_0}}x_1 + \tilde{n}$$

where $\tilde{n}$ is a zero mean complex Gaussian noise with variance of $(\frac{N_0}{2} + \frac{N_0|h_{RD}|^2E_s}{2(E_S|h_{SR}|^2+N_0)})$ per dimension[4]. The destination combines $y_{D,1}$ and $y_{D,2}^{AF}$ using maximal ratio combination rule before decoding.

Let

$$C(SNR) \triangleq \frac{1}{2}log_2(1 + SNR)\ bit/s/Hz \tag{3}$$

denote the capacity of a single Gaussian channel with signal to noise ratio SNR. The factor $\frac{1}{2}$ is introduced to account for the fact that two consecutive time slots are used for each package.

For the AF mode, it is easy to see that the achievable (instantaneous) information rate is upper bounded by the (instantaneous) mutual information of the compound channel [2]:

$$R^{AF} \leq I^{AF} = C(\|\gamma\|^2) \tag{4}$$

where

$$\gamma = \Big[\underbrace{\frac{\sqrt{E_s}|h_{SD}|}{\sqrt{N_0}}}_{\text{1st time slot}}, \underbrace{\frac{E_s|h_{RD}h_{SR}|}{\sqrt{(E_S|h_{RD}|^2+E_S|h_{SR}|^2+N_0)N_0}}}_{\text{2nd time slot}}\Big] \tag{5}$$

That h_{SR} and h_{RD} are interchangeable in the above SNR formulation suggests a capacity symmetry with respect to the position of the source and the destination in the AF mode.

2) DF Mode: In the DF mode, the relay demodulates and decodes the packet and forwards *part or all* of the information possibly using a different (compression or error control) code. We note that the DF mode we consider here has certain flavor of the *compression-forward* (CF) mode discussed in [8]). The difference, however, is that the relay in CF need not decode the message and, rather, forwards compressed versions of its observations. From network information theory, one realizes that the achievable information rate of the DF mode is determined by the max-flow min-cut of the system. The cut set around the source forms a broadcast channel and the cut set around the destination forms a parallel channel (due to the orthogonality in time). The system capacity is the minimum value of the two cut sets' capacity.

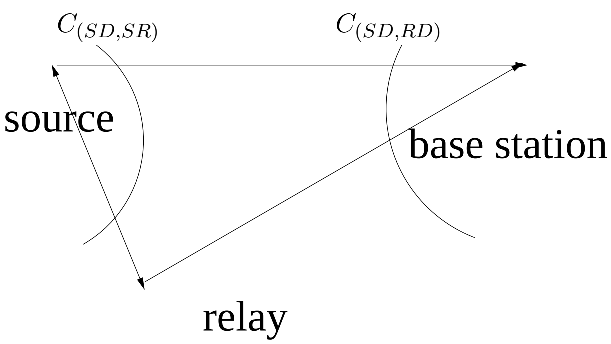

Fig. 1. Cut sets of relay system

When the instantaneous channel information is available, the channels can be treated as Gaussian channels. Since Gaussian broadcast channel is a degrade broadcast channel [7], the better channel can always decode the information intended for the worse channel also. The capacity of the broadcast cut set is the one with better signal-to-noise ratio,

$$\begin{aligned} C_{cut1} &= \max\{C_{SD}, C_{SR}\} \\ C_{SD} &\triangleq C(\frac{E_s}{N_0}|h_{SD}|^2) \\ C_{SR} &\triangleq C(\frac{E_s}{N_0}|h_{SR}|^2) \end{aligned}$$

The capacity of the cut set around the destination is the sum rate of the two paralleled channels.

$$\begin{aligned} C_{cut2} &= C_{SD} + C_{RD} \\ C_{RD} &\triangleq C(\frac{E_s}{N_0}|h_{RD}|^2) \end{aligned}$$

The system's instantaneous achievable rate is

$$R^{DF} \leq \min\{C_{cut1}, C_{cut2}\}$$

Based on the relative value of C_{SD}, C_{SR} and C_{RD}, the detailed (instantaneous) information rate for the DF mode is upper bounded by

$$R^{DF} \leq \begin{cases} C_{SD}, & \text{if } C_{SR} \leq C_{SD}, \\ C_{SR}, & \text{if } C_{SD} < C_{SR} \leq C_{SD} + C_{RD}, \\ C_{SD} + C_{RD}, & \text{if } C_{SD} + C_{RD} < C_{SR}. \end{cases} \quad (6)$$

Now that we have the instantaneous rates of the AF and DF system in (4) and (6) respectively, we can average them over the distribution of the fading coefficient (a Rayleigh distribution whose mean is some power the distance) to account for the signal attenuations caused by the channel fading and the geometry of the terminals. These results are plotted in Figures 3, 5, 8 and are discussed in the succeeding sections.

III. Capacity at Different Locations

A. Ergodic Capacity

Ergodic capacity, more commonly known as the Shannon limit, determines the maximum achievable information rate averaged over all fading states. By averaging the above mutual information results on the distribution of the channel gains, ergodic capacities can be obtained for both modes. We evaluate two cases and call them "parallel case" and "ellipse case" respectively.

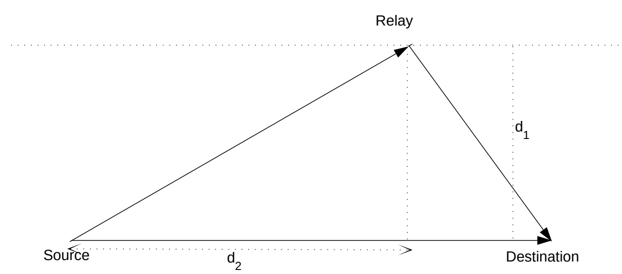

Fig. 2. Parallel case

Parallel Topology: As shown in Figure. 2, in parallel case the relay will be at various places on the horizontal dashed line. The horizontal dashed line is paralleled to the link between the source and the destination with a distance of d_1. The distance between the source and the destination is normalized to unity.

Figure.3 shows the ergodic capacity in a parallel case where $l_{SD} = 1$, $E_s = 1$, $N_0 = 1$ and $d_1 = 0.5$ (see Figure. 2). Solid lines represent the DF mode, dashed lines represent the AF mode, and power law propagation models of $\delta = 2, 3, 4$ are evaluated. From the curves, we can see that regardless of the value of δ, the capacity of the AF system exhibits a symmetry property, and the maximum value is achieved at the mid-point. The former confirms that the positions of the source and the destination are interchangeable, as is implied in (5).

To analyze the effect of the relay's location, we take the parallel case with $\delta = 2$ as an example. First, note that the effective SNR of the AF system is the sum SNRs of two spatially independent channels: the direct channel between the source and the destination, and the cascade channel consisting of the source-relay link and the relay-destination link. The SNR of the former is irrespective to the relay, and the SNR of the latter varies with the location of the relay. The average SNR of the cascade channel can be transformed into the following expression:

$$\begin{aligned} &SNR_{cas} \\ \propto&\frac{1}{l_{SR}^2 + l_{RD}^2 + l_{SR}^2 l_{RD}^2 \frac{N_0}{E_s}} \\ =&\frac{1}{\frac{N_0}{E_s}\left(l_{SR}^2 + \frac{E_s}{N_0}\right)\left(l_{RD}^2 + \frac{E_s}{N_0}\right) - \frac{E_s}{N_0}} \\ =&\frac{1}{\frac{N_0}{E_s}\left(d_2^2 + d_1^2 + \frac{E_s}{N_0}\right)\left[(l_{SD} - d_2)^2 + d_1^2 + \frac{E_s}{N_0}\right] - \frac{E_s}{N_0}} \end{aligned}$$

Where l_{SD}, l_{RD} and l_{SR} are the distances between the respective nodes. The main part of the denominator can be written as

$$\begin{aligned} &\left(d_2^2 + d_1^2 + \frac{E_s}{N_0}\right)\left[(l_{SD} - d_2)^2 + d_1^2 + \frac{E_s}{N_0}\right] \\ =&(d_2^2 + A^2)\left[(l_{SD} - d_2)^2 + A^2)\right] \\ =&l_{SD}^2 A^2 + \left[\frac{1}{4}(l_{SD}^2 - 4A^2) - \left(\frac{1}{2}l_{SD} - d_2\right)^2\right]^2 \end{aligned}$$

where $A^2 = d_1^2 + \frac{E_s}{N_0}$ is a constant value. When $l_{SD} \leq 4A^2$, the denominator achieves minimum when $d_2 = 0.5l_{SD}$. Otherwise, the denominator has two minimum values achieved when $d_2 = 0.5l_{SD} \pm \sqrt{0.25l_{SD}^2 - A^2}$. Accordingly, the SNR of the cascade channel, and subsequently the effective SNR of the AF system, will be maximized. This suggests that in the AF mode, one should look for relays that are half way between the source and the destination when the transmit power is high, or choose the nodes near the source/ destination when the transmit power is low based on l_{SD} (which is less likely to happen in real system). Our numerical results have proved this, although they are not shown here.

The same symmetry property, however, is not observable in the DF mode, The capacity of the DF mode peaks out when the relay sits at some position near the source; but unlike one would expect from a 2-by-1 MISO system, the optimal location is not that close to the source, but appears to be between the $1:9$ and $3:7$ sections from the source to the destination for different propagation models. Recall that

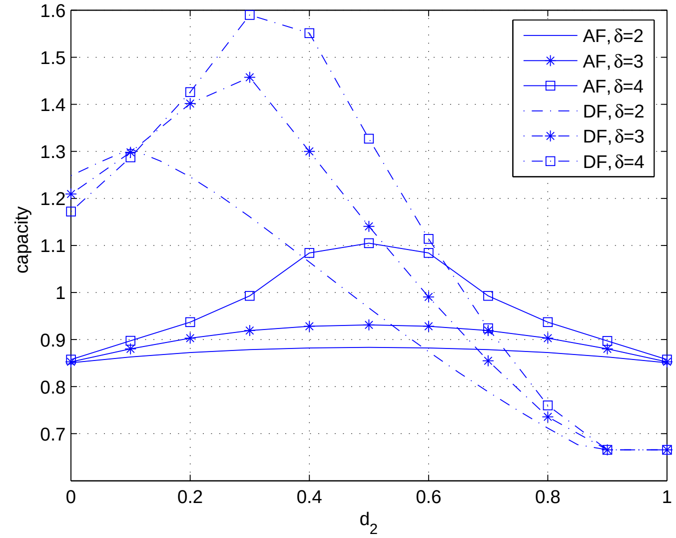

Fig. 3. Parallel case, ergodic capacity vs the location of the relay. $E_s = 1, d_1 = 0.5, l_{SD} = 1$

the capacity of DF is $\min\{C_{cut1}, C_{cut2}\}$. As the relay moves between the source and the destination, the increase of C_{cut1} will cause the decrease of C_{cut2} and vice versa. Hence, the maximum value is achieved when $C_{cut1} = C_{cut2}$. When δ is large, i.e. high order signal attenuation, C_{SR} and C_{RD} tend to dominate C_{cut1} and C_{cut2}, making the capacity curve closer to symmetric and the optimal relay location closer to the median line. Additionally, we see that each capacity curve consists of three segments. The first segment, spanning from the source to the optimal relay location, represents the case when the cut set around the destination (C_{cut2}) is the bottleneck for information flow; hence, the capacity increases as the relay moves toward the destination. The second segment represents the case when the cut set around the source (C_{cut1}) dominates, and consequently the capacity decreases as the relay moves away from the source. Finally the capacity reaches a floor that is irrelevant to the relay location. This happens when the quality of the source-relay channel is worse than the source-destination channel, i.e. $l_{SR} > l_{SD}$, and the relay system reverts to the non-cooperative mode.

Ellipse Topology: In the ellipse case, the source and the destination locate at the two foci separated by a unit distance, and the relay moves along the locus of the ellipse. The total distance from the source to the relay and finally to the destination is a constant value equivalent to the major axis of the ellipse.

Figure. 5 shows the ergodic capacity in the ellipse case where the major axis equals 1.1, i.e. $l_{SD} = 1$, $l_{SR} + l_{RD} = 1.1$. The x-axis denotes the exact distance between the source and the relay l_{SR}. The observations are similar to

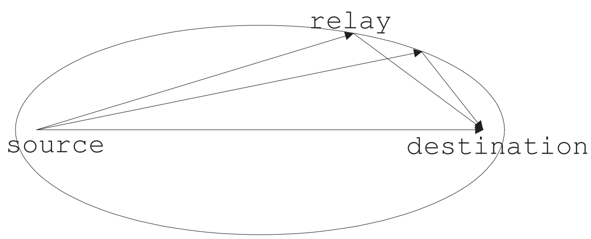

Fig. 4. Ellipse case

the parallel case: that is, the capacity curves look symmetric with respect to the node locations for the AF mode and asymmetric for the DF mode, and the optimal relay location sits at the mid-point between the source and the destination for the AF mode when the value of transmission power is not too low base on l_{SD} and closer to the source for the DF mode. We also observe that AF and DF outperform each other at different relay locations, which agrees with the results in [4].

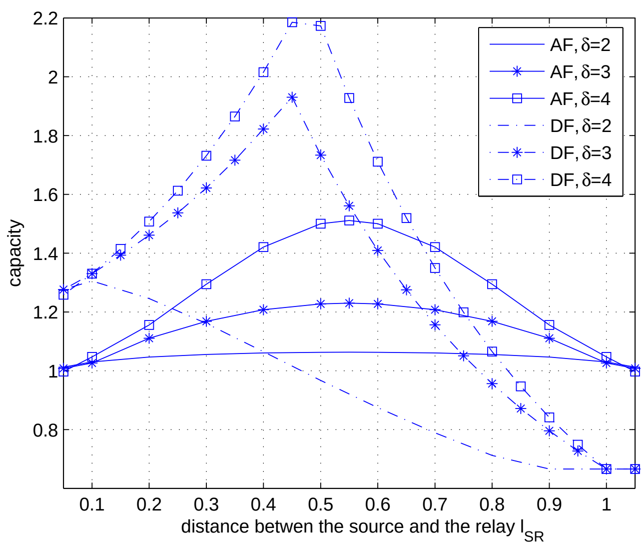

Fig. 5. Ellipse case, ergodic capacity vs the location of the relay, $E_s = 1$.

B. Outage probability

Outage capacity, aka outage probability or simply, outage, is another important statistical measure for the quality of a fading channel especially in slow fading cases. Outage specifies the probability that the instantaneous channel quality fails to meet a satisfactory threshold θ. Using information rate as the measure for channel quality, the outage probability for a single channel can be computed using:

$$P_{\text{out}}(\theta) \triangleq \Pr(R < \theta) = \int_0^\theta f_c(x)dx \tag{7}$$

where $f_c(x)$ is the probability density function of the instantaneous information rate of that channel.

As shown in (5), the AF system can be viewed as a single channel with an effective SNR γ; hence, the outage probability can be evaluated numerically using (7).

For the DF mode, the outage needs to be evaluated with respect to the cases when the cut set around the source or around the destination dominates. The former is a degraded broadcast channel problem, and outage happens when $C_{cut1} = \max(C_{SR}, C_{SD}) < \theta$. The latter is a parallel channel problem, and outage happens when $C_{cut2} = C_{SD} + C_{RD} < \theta$. Overall the outage for the DF system can be computed as

$$\begin{aligned} P_{out} &= \Pr(\max(C_{SR}, C_{SD}) < \theta) \cdot \\ &\quad \Pr\big(\max(C_{SR}, C_{SD}) < C_{SD} + C_{RD}\big) \\ &\quad + \Pr(C_{SD} + C_{RD} < \theta) \cdot \\ &\quad \Pr\big(\max(C_{SR}, C_{SD}) > C_{SD} + C_{RD}\big) \\ &= \Pr(C_{SR} < \theta)\Pr(C_{SD} < \theta) \\ &\quad + (1 - \Pr(C_{SR} < \theta))\Pr(C_{SD} + C_{RD} < \theta) \end{aligned}$$

Figures 6 and 7 show the outage probability in two topologies with threshold $\theta = 0.35$ bits per channel use when $E_s = 1$. The outage results appear quite consistent with the capacity results. For the AF mode, the outage curve is also symmetric and the lowest outage is achieved when the relay resides in equal distances between the source and the destination when transmission power is high. When transmission power is too low the optimal location is not at the midpoint(The numerical results are not shown here). For the DF mode, the optimal relay position in terms of the least outage is somewhere between the $3:7$ and $4:6$ sections from the source to the destination.

C. Capacity contour

To cast a complete view of how the system capacity relates to the geometry of the terminals and to provide engineering guidelines for choosing the optimal relay location, we plot in Figure. 8 the capacity contour for the DF mode when the relay is at different positions. To ease analysis, the source and the destination are placed at positions $(0, 0)$ and $(0, 1)$ with a normalized distance of 1. We take the case when the signal follows the cubit law attenuation ($\delta = 3$) and $E_s = 1$. It is interesting to observe that the contour curves are completed by two sets of arcs, co-centered at the source and the destination terminals, respectively. Clearly, these arcs correspond to the capacities of the two cut sets around the source and the destination. We see that the capacity is maximized by choosing a relay that sits at the $4:6$ section on the straight line between the source and the destination. The capacity starts to drop as the relay moves away from the optimal location in either direction, but at

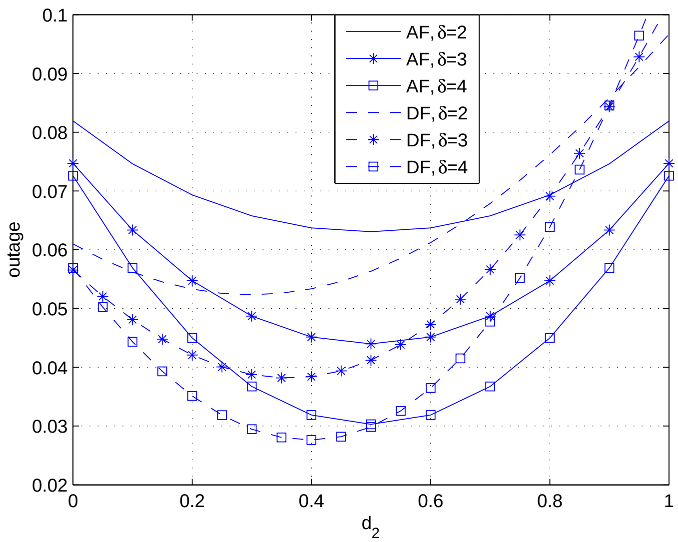

Fig. 6. Parallel case, outage probability vs the location of the relay. $E_s = 1, d_1 = 0.5, l_{SD} = 1$

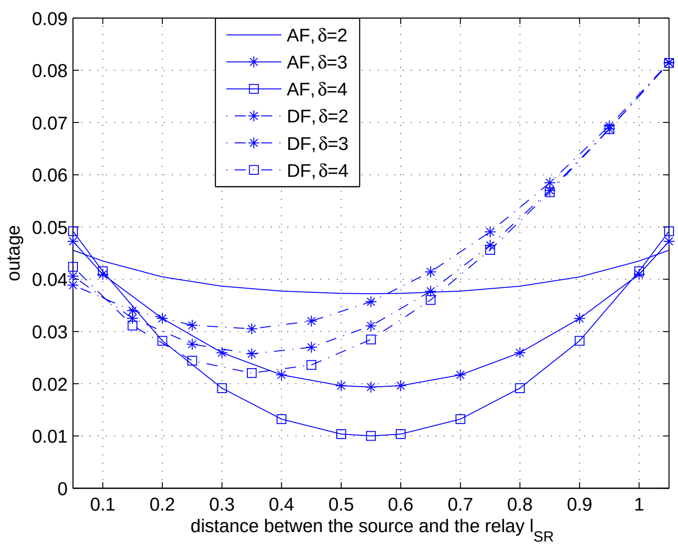

Fig. 7. Ellipse case, outage probability vs the location of the relay, $E_s = 1$

a speed that is faster toward the destination than toward the source. Finally, when the relay moves farther beyond the destination, the capacity of the direct link between the source to the destination will have exceeded that of the inter-user channel, i.e. $C_{SR} < C_{SD}$. Hence the relay node will stop message forwarding, and the cooperative system degenerates to a single-channel system with a capacity of C_{SD}.

The relation between the cooperative systems, i.e. virtual antenna arrays, with the true multi-antenna MIMO systems has been the interest for a while. Work of Kramer, Gastpar and Gupta [9] and Xie and Kumar [10] reveals that decode-forward is akin to multi-antenna transmission and compress-forward is akin to multi-antenna reception. It has further been suggested that DF will achieve the maximal capacity when the relay moves toward the source, and CF will achieve the maximal capacity when the relay moves toward the destination. Our results about DF are quite consistent with the multi-antenna interpretation. However, using practical signal attenuation models, we have found that the optimal location for the relay need not be extremely close to the source.

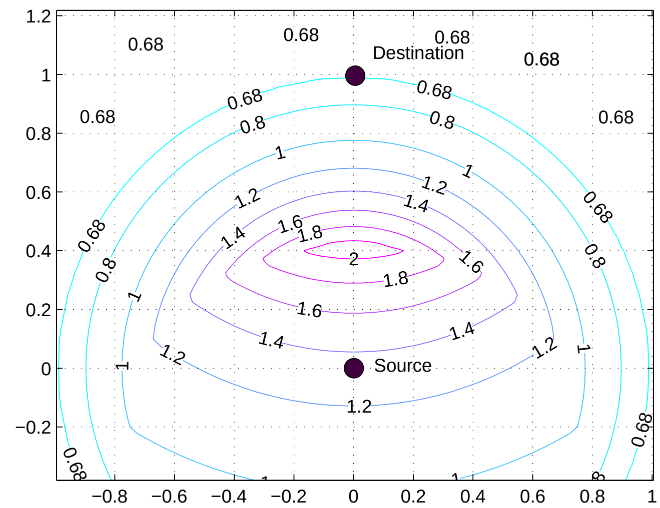

Fig. 8. Ellipse case, ergodic capacity vs the location of the relay.

To see why the DF relay system does not perform nearly as well as a 1-by-2 single-input multi-output (SIMO) system even when the relay gets very close to the destination (i.e making the relay-destination channel near-perfect), consider the difference in the decoding strategies. In the SIMO system, the signals received by the multiple antennas are optimally combined and *jointly* decoded; whereas in the DF relay system, the signals received by the virtual antenna array are *separately* decoded (i.e. the relay demodulates and decodes its received signals, and passes hard-decisions to the destination). In this sense, compression-forward appears to be the dual of decode-forward. If the compression of the received (analog) signals at the relay is near-lossless, then the destination will attain undistorted copies of all the signals received at the virtual antenna array, and can therefore perform optimal combining and joint decoding.

IV. Conclusion

We have analyzed the performances of amplify-forward and decode-forward, the two basic signal relaying strategies, for a three-terminal cooperative system in Rayleigh fading environment. The max-flow min-cut theory is used as the base approach, and the performance measure is quantified by the ergodic capacity and the outage probability. We have explicitly taken into account the geometry of the nodes, the distances between them, and the resulting attenuation of radio signals, and weighted the information as a function of the transmission distances. For the AF mode, we have demonstrated an interesting symmetry property in both the capacity and the outage results, and shown that the peak value is achieved when the relay sits half-way between the source and the destination in most cases with the exception when the power is too low. For the DF mode, our results confirm that the system operates much like a multi-antenna transmission system [10]. Using practical signal attenuation models, we found that the optimal relay location is somewhere around, but not extremely close to, the source.

References

[1] A. Sendonaris, E. Erkip and B. Aazhang, "User cooperation diversity-Part I:system description," *IEEE Trans. on Commun.*, pp. 1927-1938, Nov 2003.

[2] N. Lanenan, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: efficient protocols and outage behavior," *IEEE Trans. on info. Theory.*

[3] M. Janani, A. Hedayat, T. Hunter, and A. Nosratinia, "Coded cooperation in wireless communications: space-time transmission and iterative decoding," *Signal Processing, IEEE Transactions on*, Feb 2004.

[4] R.U.Nabar, H.Bolcskei and F.W. Kneubuhler, "Fading relay channels:performance limits and space-Time signal design,"*IEEE Journal on Selected Areas in Communications*, June 2004

[5] M. Yu and J. Li, "Is amplify-and-forward practically better than decode-and-forward or vice versa?" *Conf. ICASSP*, 2005

[6] Zinan Lin, Elza Erkip and Andrej Stefanov, "Cooperative regions for coded cooperative systems," *GLOBECOM Conf.* 2004

[7] T.Cover and J Thomas, "Elements of Information Theory," *John Wiley & Sons, Inc*

[8] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inform. Theory*, vol. 25, pp. 572-584, Sept. 1979

[9] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," submitted to *IEEE Trans. Inform. Theory*.

[10] L.-L. Xie and P. R. Kumar, "An achievable rate for the multiple level relay channel," submitted to *IEEE Trans. Inform. Theory*.