In *Proc. Allerton Conf. Commun., Contr., Computing,* (Illinois), Oct. 2000. (invited paper)

Exploiting Distributed Spatial Diversity in Wireless Networks

J. Nicholas Laneman   Gregory W. Wornell

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139
{jnl,gww}@allegro.mit.edu

Abstract

We develop energy-ef£cient transmission protocols for wireless networks that exploit spatial diversity available at distributed antennas to combat multipath fading through coordinated transmission and/or processing by several radios. We discuss the problem in a general network setting and focus on a multiple-access case with suf£cient symmetry to make the presentation concise. In particular, we examine several possibilities for the strategy employed by the assisting radios, or relays, including decoding and forwarding and amplifying and forwarding. To characterize performance, we develop outage regions and associated outage probabilities that indicate robustness of the transmissions to varying signal-to-noise ratios (SNRs). The outage regions are treated as set-functions of the observed SNR γ on the channel between the two source radios. The outage regions are thus (conditional) events de£ned in terms of the other SNR parameters of the channel, and can be readily interpreted in coded and uncoded settings. Our results suggest that the relays employ a threshold rule, namely, for SNRs satisfying $\gamma > \gamma^*$, for some value γ^*, the relays decide to cooperate and pass along each other's transmissions, preferably by decoding and re-encoding with a separate codebook for suf£ciently high γ; for SNRs satisfying $\gamma < \gamma^*$, the relays can do better by simply retransmitting their own signals.

1 Introduction

Relaying information over several point-to-point communication links is a basic building block of communication networks. Such relaying is utilized in wired and wireless networks to achieve higher network connectivity (broader coverage), ef£cient utilization of resources such as power and bandwidth, better economies of scale in the cost of long-haul transmissions (through traf£c aggregation), interoperability among networks, and more easily manageable, hierarchical network architectures.

In wireless networks, direct transmission between widely separated radios can be very expensive in terms of transmitted power required for reliable communication. High-power transmissions lead to faster battery drain (shorter network life) as well as increased interference at nearby radios. As alternatives to direct transmission, there are two basic and frequently-employed examples of relayed transmission for wireless networks. In cellular settings, for

This work has been supported in part by ARL Federated Labs under Cooperative Agreement No. DAAL01-96-2-0002, and by NSF under Grant No. CCR-9979363.

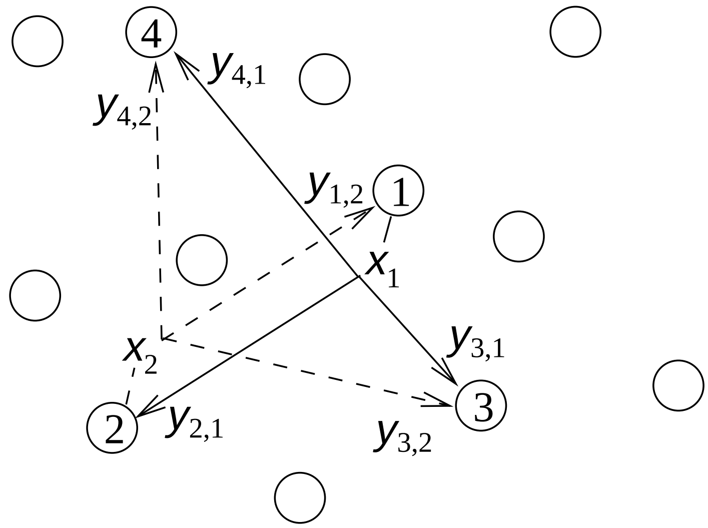

Figure 1: Example wireless network in which transmission protocols for exploiting distributed spatial diversity can be motivated. Indicated are transmitted signals x_i and received signals $y_{j,i}$. Throughout the paper, we focus on transmissions from radios 1 and 2 to radios 3 and 4, respectively.

example, networks provide connectivity between low-power mobiles by providing local connections to high-power basestations that are relayed via a wireline basestation network. In sensor networks, and military battle£eld communication networks in general, the use of wireline infrastructure is often precluded and the radios may be substantially power constrained; for these ad-hoc or peer-to-peer networks, transmissions can be relayed wirelessly. As these examples suggest, relayed transmission enlists two or more radios to perform multiple transmissions. The end-to-end transmissions potentially incur higher delay, but because the individual transmissions are over shorter distances (in the wireless case), or over high-quality cabling (in the wireline case), the power requirements for reliable communication can be much lower.

The basic relaying protocols described above are constructed from the sequential use of point-to-point links, where the links are essentially viewed at the network protocol layer; however, more general approaches are possible that involve the coordination of *both* the direct and relayed transmissions, at the network and lower protocol layers, and correspond to scenarios to which the classical relay channel model applies. (See [1] and related work in [2].) In this paper, we develop energy-ef£cient relaying protocols that exploit spatial diversity available at distributed antenna elements to combat fading due to multipath propagation, a particularly severe form of interference experienced in wireless networks.

To illustrate the main concepts, we consider the wireless network depicted in Fig. 1. At the physical layer, destination radios receive potentially useful signals from all transmitters that are active, and may combine multiple transmissions of the same signal to reduce variations in performance caused by signal fading, a technique referred to broadly as spatial diversity combining [3]. We refer to this form of spatial diversity as *distributed spatial diversity*, in contrast to the currently more conventional forms of spatial diversity [4], because the radios essentially share their antennas and other resources to create a "virtual array" through distributed transmission and signal processing.

After developing a mathematical model in Section 2 for the network in Fig. 1, we scratch the surface of the rich set of design issues and options that arise in the context of exploiting distributed spatial diversity for wireless networks. Section 3 casts the basic relaying protocols, referred to as direct and multihop transmission, respectively, into our framework, and explores a number of possibilities for diversity transmission and hybrid protocols, in terms of what signals the source and relay jointly transmit as well as how the relay and destination jointly process signals. Section 4 develops outage regions for the various transmission options, and Section 5 compares these regions and their corresponding outage probabilities. Performance comparisons in Section 5 suggest that our diversity transmission protocols are capable of overcoming the noisy channels between the distributed radio antennas to achieve diversity gain and outperform direct and multihop transmission in a variety of scenarios of interest.

2 System Model

Our network model consists a of collection of M radios that share L orthogonal channels. A transmission period consists of two consecutive blocks, and the channels are allocated to (up to) L radios during each transmission period. During the £rst block of a transmission period, a radio transmits on its assigned channel, and receives on a separate channel. Radios might choose the strongest channel or might select a channel at random for reception. Such options complicate the story for $L > 2$, so we consider the case of $L = 2$ for simplicity of exposition. Depending upon the strength of the signal received on the selected channel, the radio decides between resending its own transmission (or, more generally, additional parity bits from a more ef£cient code, *e.g.*, rate-compatible punctured codes) in the next block, or relaying the other radio's received signal in the next block. As a result, equal bandwidth and power allocations seem to be a natural choice.

In our model for the wireless network depicted in Fig. 1, narrowband transmissions suffer the effects of path loss and ¤at fading as arise in *e.g.*, slow-frequency-hop networks. Our analysis focuses on the case of slow fading to isolate the bene£ts of spatial diversity alone; however, we emphasize at the outset that our results extend naturally to the kinds of highly mobile scenarios in which faster fading is encountered.

Our baseband-equivalent discrete-time channel model for the network consists of two subchannels, orthogonal in, *e.g.*, adjacent frequencies. This decomposition is necessary because practical limitations in radio implementation prevent the relays from simultaneously transmitting and receiving on the same channel. Thus, radio 1 transmits on channel 1 and receives on channel 2 to potentially relay the signal transmitted by radio 2, and vice versa. The received signals at all four radios are modeled by

$$y_{j,i}[n] = a_{i,j}\, x_i[n] + z_{j,i}, \quad i = 1,2, \quad j = 1,2,3,4, \quad j \neq i. \tag{1}$$

Here $a_{i,j}$ captures the effects of path loss and static fading on transmissions from radio i to radio j, x_i is the transmitted signal of radio i having average energy $\mathcal{E}_i$, and $z_{j,i}[n]$ models additive receiver noise and other forms of interference at receiver j in channel i.

Statistically, we model the fading coef£cients $a_{i,j}$ as zero-mean, mutually independent complex random variables with variances $\sigma^2_{a_{i,j}}$, and we model the additive noises $z_{j,i}[n]$ as zero-mean, mutually independent, white complex jointly Gaussian sequences with variance $\mathcal{N}_j$. The network geometry is assumed unknown or too dynamic to track, so the channels may be well modeled as having i.i.d. signal-to-noise ratios (SNRs) under an appropriate distribution. We denote the SNR in each received signal as $\gamma_{i,j} \stackrel{\triangle}{=} |a_{i,j}|^2\, \mathcal{E}_i/\mathcal{N}_j$. For example, under the Rayleigh fading model, the SNRs are independent exponential random variables.

As we develop our transmission protocol in Section 3, it will be convenient to consider successive pairs of data blocks from the channel model in (1). Speci£cally, for blocklength N, we collect the appropriate samples into the vectors

$$\begin{aligned} \mathbf{x}_i[k] &= \left[x_i[kN] \quad x_i[kN+1] \quad \cdots \quad x_i[kN+(N-1)]\right]^{\mathrm{T}}, \\ \mathbf{y}_{j,i}[k] &= \left[y_{j,i}[kN] \quad y_{j,i}[kN+1] \quad \cdots \quad y_{j,i}[kN+(N-1)]\right]^{\mathrm{T}}, \\ \mathbf{z}_{j,i}[k] &= \left[z_{j,i}[kN] \quad z_{j,i}[kN+1] \quad \cdots \quad z_{j,i}[kN+(N-1)]\right]^{\mathrm{T}}. \end{aligned} \tag{2}$$

3 Transmission Protocol

Throughout this paper, we focus on relatively simple protocols that operate on two consecutive blocks indexed by $2k$ (even) and $2k+1$ (odd). At a high level, our protocols involve the following steps:

- **Even Blocks:** Radios $i = 1, 2$ encode new information into blocks $\mathbf{x}_i[2k]$, respectively. Radios $j = 1, 2, 3, 4$ receive signals $\mathbf{y}_{j,i}[2k]$, $i = 1, 2$. Radios 3 and 4 defer their processing until the end of block $2k+1$. The transmitting radios process their respective received signals and decide whether they will cooperate in the next block, and if so, how.

- **Odd Blocks:** Radios $i = 1, 2$ encode either their own or their partner's data into blocks $\mathbf{x}_i[2k+1]$, respectively. Radio 3 receives signals $\mathbf{y}_{3,i}[2k+1]$, $i = 1, 2$, and jointly processes these signals with $\mathbf{y}_{3,i}[2k]$, $i = 1, 2$, received in the previous block. Radio 4 operates on its respective received signals in similar fashion.

Among many possible coordination strategies, we consider a simple protocol in which the two cooperating radios accurately estimate the SNR $\gamma_{2,1} = \gamma_{1,2}$ between them and use this estimate to select a suitable cooperative action. Such a protocol allows the radios to retransmit (in the form of repetition codes or more sophisticated single-user coding schemes such as rate-compatible punctured codes) their own information when $\gamma_{1,2}$ is too small to justify cooperation, and to transmit each other's information (again, in the form of repetition codes or more powerful joint encoding schemes) when $\gamma_{1,2}$ is large enough to justify cooperation and, in particular, to provide spatial diversity bene£t to the transmitting radios. In any case, a destination radio can appropriately combine its received signals by exploiting control information in the protocol headers, *e.g.*, a £eld indicating the cooperative action taken by the transmitting radios. While the nature and amount of this control information, as well as the accuracy and consistency of the SNR estimates at the two cooperating radios, are important practical considerations, a detailed study of their impact on system performance is beyond the scope of this paper.

Assuming both radios estimate the realized value $\gamma_{1,2}$ perfectly, they will choose identical cooperation strategies due to the statistical symmetry of the channels implied by our model. We examine three options for the transmissions in (odd) block $2k+1$:

- **Direct Transmission:** The coordinating radios ignore each other's transmissions and re-transmit their information from block $2k$.

- **Multihop Transmission:** The coordinating radios fully decode and retransmit each other's information, and the destinations ignore the initial transmissions from block $2k$, processing only the relayed transmissions in block $2k+1$.

- **Diversity Transmission:** The coordinating radios assist one another by transmitting each other's information in block $2k+1$. Reasonable strategies explored in our previous work include decoding and forwarding as well as amplifying and forwarding.
  - **Decode and Forward:** If $\gamma_{1,2}$ is reasonably large, it is advantageous for the relay to decode the transmissions to suppress noise on the channel between the source and relay, and re-encode the signal, potentially with a different codebook, for transmission between the relay and destination.
  - **Amplify and Forward:** For situations in which $\gamma_{1,2}$ is small, a linear relay that simply ampli£es its received signal can be shown to be more effective than decoding and forwarding.

To summarize, each protocol consists of a source codebook, a relay processing/coding scheme, and a destination decoder.

Our combined analysis and empirical studies suggest that we may employ a pair of threshold tests on the SNR between the cooperating radios to choose the strategy with best expected performance, as measure by, *e.g.*, the lowest (conditional) outage probability. To develop this result, we examine the outage regions associated with each case in Section 4, and compare conditional outage probabilities in Section 5. We stress at the outset that none of the protocols we propose are necessarily optimal, but they represent reasonable protocols whose performance we can evaluate and begin to optimize. Improving upon these protocols, and developing others, is the subject of on-going work.

4 Outage Regions

Generalizing upon our previous results for uncoded, *i.e.*, $N = 1$, single-user systems with a probability of bit-error performance measure [5], we characterize the performance of the various cooperation strategies in terms of outage regions and outage probabilities. Outage regions speci£ed in terms of the SNR random variables $\gamma_{i,j}$ have convenient interpretations in both coded and uncoded settings, but we will develop our results from a coded perspective and determine events in which the realized capacity of the channel falls below a target transmission rate. We convert this event into an event de£ned in terms of the SNRs in the channel.

Since the capacity is a function of the SNR random variables of the channel, it too is a random variable. The event that this capacity random variable falls below some £xed rate R is referred to as an *outage event* (or *outage region* in terms of the SNR random variables), because reliable communication is not possible inside this region. The probability of an outage event is referred to as the *outage probability* of the channel,

$$P_{\text{out}}(R) = \Pr\left[\mathsf{C} < R\right]. \tag{3}$$

We stress that outage regions are independent of the distribution of the underlying random variables, while outage probabilities are intimately tied to them. For example, if the outage region of a channel at a particular rate is a strict subset of the outage region of another channel at that rate, then the £rst channel has smaller outage probability regardless of the probability distribution on the channel parameters. Furthermore, as we will see, several of our cooperation strategies appear to have similar outage probabilities, but the structure of their outage regions is suf£ciently different that we might prefer one over the other in various regimes. As a result, both outage regions and outage probabilities will be useful for characterizing our transmission protocols.

We consider transmission from radio s to radio d, with the potential of relaying the transmissions of radio r (and having radio r relay the transmissions of radio s). We parameterize the results in this form for compactness, but note that they can immediately be interpreted from the perspective of radio 1 by setting $s = 1$, $r = 2$, and $d = 3$; similarly, to interpret the results from the perspective of radio 2, we set $s = 2$, $r = 1$, and $d = 4$.

In the following sections, we determine the outage regions for direct transmission, multihop transmission, diversity transmission with amplifying and forwarding, and diversity transmission with decoding and forwarding.

4.1 Direct Transmission

Direct transmission in our setting corresponds to a point-to-point communications channel, to which we may readily apply classic information theoretic arguments [1]. Speci£cally, the capacity between the source and destination radios using repetition coding satis£es

$$C_{\text{SH}} = C_{s,d} = \frac{1}{2}\log\left(1 + 2\gamma_{s,d}\right), \tag{4}$$

with $\mathbf{x}_s[2k] = \mathbf{x}_s[2k+1]$ (repetition coding) distributed as i.i.d. zero-mean complex Gaussian random variables each with variance $\mathcal{E}_s$. We note that, while in principle more powerful forms of coding than repetition are possible across the two blocks $\mathbf{x}_s[2k]$ and $\mathbf{x}_s[2k+1]$, comparison to diversity transmission with amplifying and forwarding, inherently analogous to repetition, is most convenient in the repetition coded case.

Inspecting (4), we see that the outage event $C_{\text{SH}} < R$ can be readily manipulated into an event de£ned in terms of the SNR random variable between the source and destination, *i.e.*,

$$\gamma_{s,d} < t/2, \tag{5}$$

for an appropriate SNR threshold t that increases with increasing R.

4.2 Multihop Transmission

Multihop transmission corresponds to direct transmission between the source and relay radios followed by direct transmission between the relay and destination radios. Thus we might expect the capacity of the cascade of the two channels to be the minimum of the capacities. Indeed, results on cascade channels [1] yield

$$C_{\text{MH}} = \min\{C_{s,r},\ C_{r,d}\} = \frac{1}{2}\min\{\log\left(1 + \gamma_{s,r}\right),\ \log(1 + \gamma_{r,d})\}, \tag{6}$$

where, again, equality is achieved for complex Gaussian signals, and where we lose the factor of 2 present in (4) because the destination only processes the signal received from the relay and ignores the signal transmitted by the source.

Again, we may determine the outage region corresponding to (6), yielding

$$\gamma_{s,r} < t \quad \text{or} \quad \gamma_{r,d} < t\ . \tag{7}$$

As we will see in Section 4.4, this outage region is a strict superset of the outage region for diversity transmission with decoding and forwarding. This observation allows us to eliminate multihop transmission from our comparisons in Section 5.

4.3 Diversity Transmission with Amplifying and Forwarding

Under diversity transmission with amplifying and forwarding, the relay scales its received sequence by

$$\beta = \sqrt{\frac{\mathcal{E}_r}{|a_{s,r}|^2\,\mathcal{E}_s + \mathcal{N}_r}}$$

to satisfy its average power constraint. Note that we allow the ampli£er gain to depend upon the fading coef£cient $a_{s,r}$ between the source and relay, which the relay estimates to high accuracy. This transmission scheme can be viewed as repetition coding from two separate transmitters, except that the transmitters may have different power levels and the relay transmitter actually ampli£es its receiver noise. Nevertheless, the channel can be viewed as a single-user Gaussian noise channel and has capacity

$$C_{DA} = \frac{1}{2}\log(1 + \gamma_{s,d} + f(\gamma_{s,r}, \gamma_{r,d}))\ , \tag{8}$$

for x_s zero-mean complex Gaussian with variance $\mathcal{E}_s$, where

$$f(x,y) = \left[x^{-1} + y^{-1} + (xy)^{-1}\right]^{-1}\ ,$$

analogous to a parallel combination of resistances with values x, y, and xy, respectively.

The outage region becomes the SNR event

$$\gamma_{s,d} + f(\gamma_{s,r}, \gamma_{r,d}) < t\ , \tag{9}$$

and an outage region "boundary" for each realization of $\gamma_{s,r}$ follows after careful manipulation of (9), obtaining

$$\gamma_{r,d} = \begin{cases} 0, & \gamma_{s,d} \geq t \\ \frac{1+\frac{1}{\gamma_{s,r}}}{\frac{1}{t-\gamma_{s,d}} - \frac{1}{\gamma_{s,r}}}, & t - \gamma_{s,r} < \gamma_{s,d} < t \\ +\infty, & \gamma_{s,d} < t - \gamma_{s,r}. \end{cases} \tag{10}$$

4.4 Diversity Transmission with Decoding and Forwarding

When the SNR between the source and relay is reasonably high, it is advantageous for the relay to decode the transmission and re-encode, potentially with a different codebook, for cooperative transmission to the destination. When the relay decodes successfully, the source and relay have the same information and can be viewed as antenna elements in a transmit diversity array that uses orthogonal channels (expands bandwidth). One can show that, under the constraint that the relay decodes perfectly, for $\mathbf{x}_r[2k+1] = \hat{\mathbf{x}}_s[2k]$, *i.e.*, the codebook is a repetition code, the capacity satis£es

$$C_{\text{DD}} = \frac{1}{2}\min\{\log\left(1 + \gamma_{s,r}\right),\ \log\left(1 + \gamma_{s,d} + \gamma_{r,d}\right)\}\ . \tag{11}$$

Since we require the relay to decode perfectly, the outage regions correspond to the SNR event

$$\gamma_{s,r} < t \quad \text{or} \quad \gamma_{s,d} + \gamma_{r,d} < t, \tag{12}$$

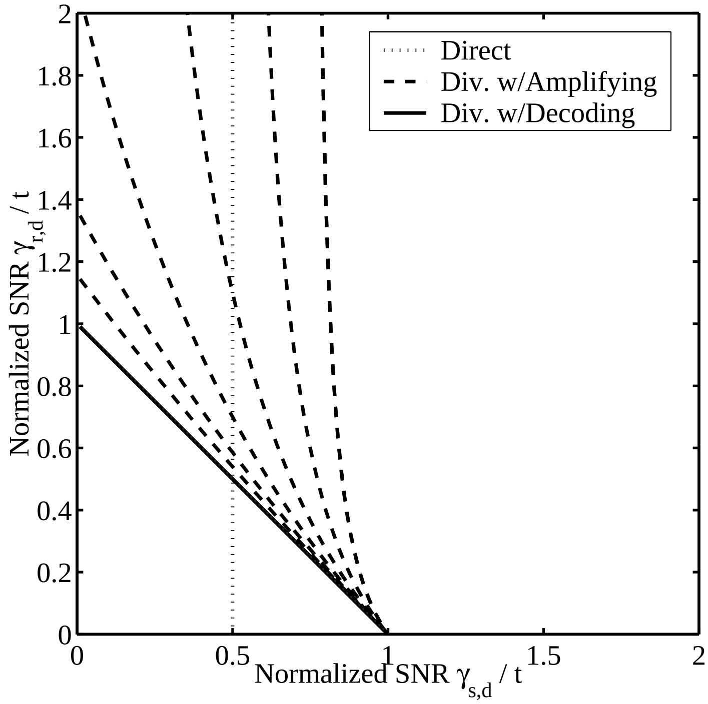

Figure 2: Outage region boundaries for repetition-code transmission protocols. The dotted line at $\gamma_{s,d}/t = 1/2$ corresponds to the outage region boundary for direct transmission. Successively lower dashed curves correspond to diversity transmission with amplifying and forwarding assuming received signal strength measurements of $\gamma_{s,r}/t = 1/4, 1/2, 1, 2, 4, 8$. The solid line corresponds to diversity transmission with decoding and forwarding for the case in which $\gamma_{s,r}/t > 1$; otherwise, the outage event for decoding is the entire plane. The outage regions lie below and/or to the left of the solid lines, *i.e.*, all include the point $(0, 0)$.

which, as we observed in Section 4.2, is a subset of the outage region for multihop transmission. For each realization of $\gamma_{s,r}$, the boundary of the outage region (12) can be manipulated into the form

$$\gamma_{r,d} = \begin{cases} 0, & \gamma_{s,r} \geq t, \quad \gamma_{s,d} \geq t \\ t - \gamma_{s,d}, & \gamma_{s,r} \geq t, \quad \gamma_{s,d} < t \\ +\infty, & \gamma_{s,r} < t, \quad 0 \leq \gamma_{s,d} < \infty \end{cases} . \tag{13}$$

5 Protocol Design and Performance

Fig. 2 depicts the outage regions for direct and diversity transmission with repetition coding, and clearly indicates improved performance (successively smaller outage regions) for amplifying and forwarding with increasing $\gamma_{s,r}$, approaching the outage region of decoding and forwarding for $\gamma_{s,r}$ large relative to the threshold t. We will see that for a suf£ciently large value of $\gamma_{s,r}$, the (conditional) outage probability for amplifying and forwarding will be smaller than that of direct transmission, even though the outage regions for direct and diversity transmission are not nested one way or the other. The same is true for comparisons between transmit antenna diversity systems (without beamforming) with one and two antennas. In practice, the dual antenna systems perform better than the single antenna systems.

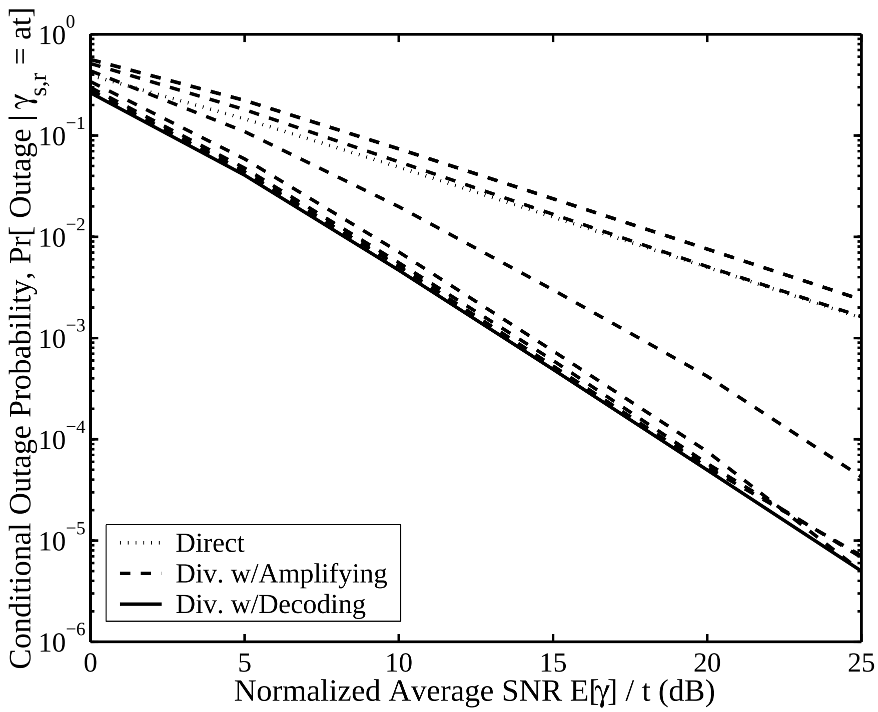

Figure 3: Conditional outage probabilities under exponential statistics. The dotted curve corresponds to direct transmission, while the solid curve corresponds to diversity transmission with decoding and forwarding for $\gamma_{s,r}/t \geq 1$. (For $\gamma_{s,r}/t < 1$, the conditional outage probability of diversity transmission with decoding and forwarding is 1.) Successively lower dashed curves correspond to diversity transmission with amplifying and forwarding for $\gamma_{s,r}/t = 1/4, 1/2, 1, 2, 4, 8$.

Indeed, conditional outage probability calculations appear to con£rm this hypothesis. For example, Fig. 3 shows outage probability calculations for the various transmission strategies, assuming the SNRs are i.i.d. exponential random variables with mean $\mathrm{E}\left[\gamma\right]$. Moreover, the results in Fig. 3 suggest thresholds for selecting among the various transmission strategies. Speci£cally, for $\gamma_{s,r}/t < 1/2$, these results suggest that direct transmission is preferable. For $1/2 \leq \gamma_{s,r}/t < 1$, bearing in mind that the conditional outage probability for diversity transmission with decoding and forwarding is 1 given $\gamma_{s,r}/t < 1$, we should employ diversity transmission with amplifying and forwarding. Finally, for $\gamma_{s,r}/t \geq 1$, the decoding relay offers uniformly lower conditional outage probability than direct transmission or diversity transmission with amplifying and forwarding.

Fig. 4 shows empirical outage probabilities for our various protocols for the case in which all the SNRs in the channel are i.i.d. exponential random variables with mean $\mathrm{E}\left[\gamma\right]$. These results suggest that energy savings on the order of 10 dB at $P_{\text{out}} = 10^{-3}$ can be obtained using protocols developed in this paper, because these protocols ef£ciently create multiple, independently-faded transmissions at separate radios that can be effectively combined at the destination receiver to achieve diversity gains. We also note that a protocol that always employs diversity transmission with amplifying and forwarding performs almost as well as our optimized hybrid protocol.

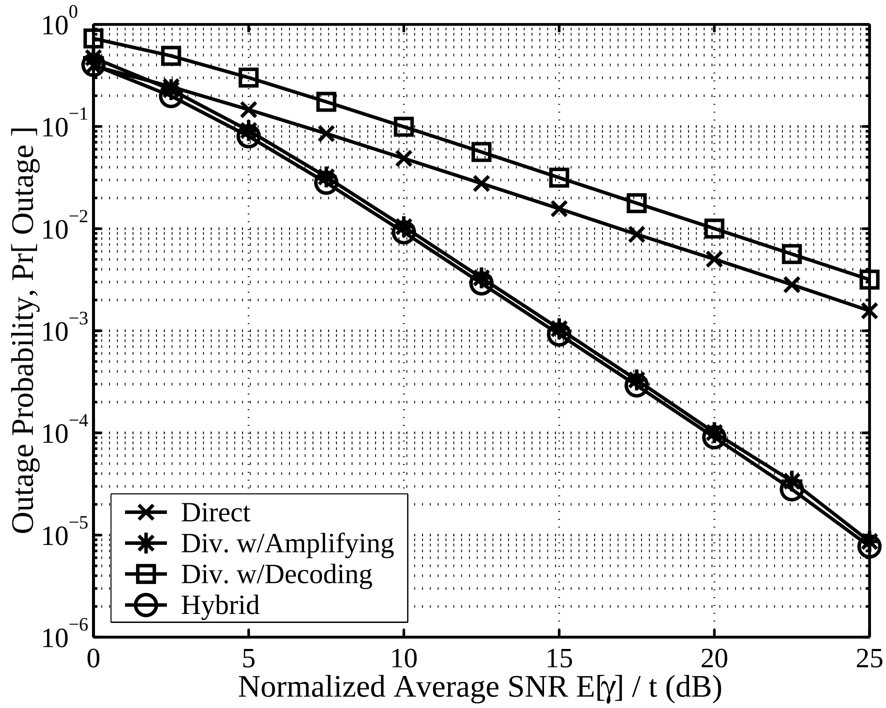

Figure 4: Outage probabilities under exponential statistics.

References

[1] Thomas M. Cover and Joy A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York, 1991.

[2] Andrew Sendonaris, Elza Erkip, and Behnaam Aazhang, “Increasing uplink capacity via user cooperation diversity,” in *Proc. IEEE Internation Symposium on Information Theory (ISIT)*, Cambridge, MA, Aug. 1998.

[3] John G. Proakis, *Digital Communications*, McGraw-Hill, Inc., New York, Third edition, 1995.

[4] Aradhana Narula, Mitchell D. Trott, and Gregory W. Wornell, “Performance limits of coded diversity methods for transmitter antenna arrays,” *IEEE Trans. Inform. Theory*, vol. 45, no. 7, pp. 2418–2433, Nov. 1999.

[5] J. Nicholas Laneman and Gregory W. Wornell, “Energy-ef£cient antenna sharing and relaying for wireless networks,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Chicago, IL, Sept. 2000.