Time-division multiplexing for green broadcasting

Pulkit Grover and Anant Sahai
Wireless Foundations, Department of EECS
University of California at Berkeley
Email: {pulkit, sahai} @ eecs.berkeley.edu

Abstract

The problem of minimizing the total (transmit and decoding) energy required for communicating over a two-receiver Gaussian broadcast channel is investigated. For achieving a specified set of rates, joint broadcast schemes (e.g. superposition coding) require smaller transmit energy per-bit than the simpler time-division multiplexing based schemes. However, for short distance communication, the energy expended in the decoding can be comparable to that required in the transmission. It is shown that in some typical short and moderate distance communication scenarios, time-division multiplexing saves on the decoding energy, thereby requiring smaller *total* energy than any joint broadcasting scheme for achieving the target rate and error probabilities. Further, it is shown that the *ratio* of the distances of the receivers from the transmitter is an important factor in the scheme design.

I. Introduction

Shannon theory has been quite successful in understanding the minimum transmit power required for achieving a specified performance in many wireless network problems. In point-to-point communication, for example, waterfall curves provide a complete characterization of the minimum transmit power required to achieve a given rate and error probability. More generally, capacity theorems for average input power limited multiuser wireless network problems can be interpreted as results on the minimum required transmit power to communicate reliably at the desired rates. For some multiuser problems of practical interest, e.g. the multiple-access channel (MAC) and the Gaussian broadcast channel, the capacity region and hence the required transmit power(s) to achieve specified rates are known exactly [1, Pg. 403–407 and Pg. 427–428]. Recent results [2], [3] have succeeded in finding the required transmit power to within a constant factor for high rates for relay networks with single source and single destination as well[1].

These results have found enormous applicability in optimizing the energy consumption in long distance wireless communiation networks. At shorter distances (which are of increasing interest), the *processing energy* is comparable to, and can even dominate the transmit energy [5], [6]. The natural problem of interest is therefore to minimize the *total* energy, that includes the transmit as well as the processing energy.

Processing energy has been addressed extensively in the literature on wireline networks[2]. However, an error free channel or a channel that drops packets, though a good model for wired networks, is not sufficiently rich to capture important aspects (e.g. noise) of a wireless communication channel. While the information theoretic perspective on the problem has mostly been limited to the point-to-point problem [5], [7], [8] (and hence oversimplifies the network structure), the obtained results have been insightful. In [8], the authors model the transmitter as a black-box that consumes energy a constant amount of energy per unit time when the transmitter is transmitting (and hence 'on'). It consumes zero energy when the transmitter is 'off'. Unlike the results from Shannon-theoretic analysis [9], the authors show that the transmissions need to be 'bursty' in that the energy minimizing rate is *non-zero*. A more refined model that depends on the system performance (transmit power, desired rate and target error probability) is introduced in [7]. This model accounts for the *decoding energy* expended at the receiver. The ensuing analysis reveals that the there is a tradeoff between the transmit and the processing energy — to minimize the total energy, transmit energy should not be studied in isolation. Further, contrary to the implication of the waterfall curve, the total energy per-bit increases unboundedly as the desired reliability increases.

To understand effects of processing energy in the design of more general networks, the authors in [10] consider the multiple-access channel (MAC). They model the processing energy at the transmitters by the black-box model that they introduced in [8]. Interestingly, the authors show that simple time-division multiple access based schemes can attain desired rates with smaller total energy than that required schemes in which the two transmitter operate simultaneously. The intuition is similar to that for the bursty strategy in the point-to-point problem: when using time-division multiplexing, one transmitter can turn itself off and save on the processing energy when the other transmitter is operating.

To investigate the impact of *decoding energy* in networks, it is more pertinent to consider the broadcast problem because it is the receivers that expend the decoding energy. We provide the system model for a Gaussian broadcast problem in Section II. Of particular interest are *joint broadcast schemes* (e.g. superposition or dirty-paper coding) that can achieve the minimum transmit power in the limit of small error probability. In Section III-A, we provide lower bounds on the required decoding complexity

[1] At high SNR, finding the capacity region within a constant number of bits can in most cases be interpreted as finding the required transmit power(s) within a constant factor. For a case when this is *not* true, see [4].

[2] We refer the reader to [7] for a survey of the related references.

(measured in number of iterations for a message passing decoder, see Section II-A) for a Gaussian broadcast system that uses a joint broadcast scheme. These complexity bounds are then used to provide lower bounds on the total energy consumed by joint broadcast schemes. In Section III-B we provide similar lower bounds to a time-division multiplexing (TDM) scheme that transmits to the two users in different time slots. At the cost of higher transmit energy, the TDM scheme saves on the decoding energy because the two users do not need to use the entire block to decode. In Section III-C we compare our bounds on the performance of the joint strategies and TDM, and conclude that at short to moderate distances (~ 1000 m or smaller at 3 GHz), TDM requires smaller total energy than joint broadcast schemes. Interestingly, it turns out that the *ratio* of the distances of the two receivers from the transmitter is an important factor in the design. In Section III-D, we show that allowing flexibility in the desired rate would not alter our results substantially.

Strictly speaking, we are comparing two *lower* bounds which may not be meaningful. However, there exist schemes that achieve within a constant factor of the lower bounds for TDM provided here in the limit of small error probability for TDM (which uses schemes for point-to-point channels). It is shown in [7], that the lower bounds for point-to-point channels take the form

$$l \gtrsim \frac{1}{\log_2(\alpha)} \log_2\left(\frac{\log_2\left(\frac{1}{\langle P_e \rangle}\right)}{K(C-R)^2}\right), \tag{1}$$

for some constant K, for R close to the channel capacity C, where α is the maximum connectivity of a node at the decoder (defined precisely in Section II-A). It is known [11] that for regular LDPCs, the error probability decays doubly exponentially in the number of iterations. Thus regular LDPCs achieve the $\log_2\left(\log_2\left(\frac{1}{\langle P_e \rangle}\right)\right)$ term within a constant factor regardless of the error probability. It is shown in [7] that the transmit power P^* that minimizes the total energy is such that $C(P^*) - R$ converges to a constant as $\langle P_e \rangle \to 0$. Thus the bound in (1) is achievable within a constant factor, and the lower bound on energy thus derived is also achievable within a constant factor. TDM essentially uses point-to-point schemes, and hence the same statement holds for achievable energy using TDM. In fact, because these bounds underestimate the decoding energy, and because it is TDM that makes more efficient use of the decoding energy, we suspect that in practice, the performance of TDM vis-a-vis joint schemes would be better than the results presented here.

II. Problem statement

A vector of length m is denoted in bold with superscript m (e.g. $\mathbf{X}^m$). A single transmitter Tx transmits to two users, user 1 and 2, across a memoryless additive white Gaussian noise channel with input X and outputs Y and Z respectively. The information transmitted to user i is a k_i length sequence of Ber(0.5) iid bits denoted by $\mathbf{B^{(i)}}^{k_i}$, $i = 1, 2$. The transmission is carried out in blocks of length m. The encoder mapping is denoted by ψ that maps $(\mathbf{B^{(1)}}^{k_1}, \mathbf{B^{(1)}}^{k_1}) \to \mathbf{X}^n$. Thus,

$$\begin{aligned} \mathbf{Y}^m &= h_1 \mathbf{X}^m + \mathbf{W}_1^m \\ \mathbf{Z}^m &= h_2 \mathbf{X}^m + \mathbf{W}_2^m, \end{aligned}$$

where h_i is the distance-dependent fade coefficient, and the elements of $\mathbf{W}_i^m$ are distributed $\mathcal{N}(0, \sigma_0^2)$, are independent over i and iid over time. We assume that the noises are thermal and hence $\sigma_0^2 = \kappa T$ (here κ is Boltzmann's constant, and the temperatures at the two receivers are assumed to be equal). We assume the decay in the signal power is according to the power law, that is, $h_i^2 = \frac{1}{4\pi (r_i/\lambda_s)^2}$, where λ_s is the wavelength of the transmitted signal. Also, $r_2 \geq r_1$, and hence user 1 (user 2) would also be referred to as the 'strong' ('weak') user respectively. The desired rates of communication are $R_i = \frac{k_i}{m}$. Since the block length is the same for the two users, R_i satisfy $\frac{k_1}{R_1} = \frac{k_2}{R_2} = m$. The objective is to achieve an average *bit-error probability* smaller than $\langle P_e^{(i)} \rangle_0$ (averaged over the channel realizations and the messages) for $i = 1, 2$. We will sometimes consider an AWGN test channel pair (G, J) where the first user has noise variance σ_G^2, and the second user has noise variance σ_J^2. The symbol 0 is reserved for the underlying channel of noise variance σ_0^2. Under a test channel pair (G, J), the average bit-error probability is denoted by $\langle P_e^{(1)} \rangle_G$ for user 1, and by $\langle P_e^{(2)} \rangle_J$ for user 2. We denote the average transmit power by P. For simplicity we assume that $m = \frac{k_1}{R_1} = \frac{k_2}{R_2}$ and hence the two messages are conveyed in the same block of length m. Probability of noises $\mathbf{W}_1^m$ taking values in a set A under a test channel G would be denoted by $\Pr_G(A)$. Similar notation is used for $\mathbf{W}_2^m$ under a test channel J and for the underlying channel. The parameter $\zeta = \frac{r_2^2}{r_1^2} = \frac{h_1^2}{h_2^2}$ is of interest as explained in Section I.

A. Decoding energy model

Our focus is on the parallelism of the decoders and the energy consumed within them. The decoding energy model is borrowed from [7] which is based on the iterative decoding model [12]. We assume that the decoder is physically made of computational nodes that pass messages to each other in parallel along physical (and hence unchanging) wires. A subset of nodes are designated 'message nodes' in that each is responsible for decoding the value of a particular message bit. Another subset of nodes, called the 'observation nodes' has members that are each initialized with at most one observation of the received channel output symbols. There may be additional computational nodes to merely help in decoding. In a departure from the model in [7] [13], we assume that the observation nodes and the message nodes are *disjoint*. This allows simplicity in our exposition, while not altering the flavor of our results.

Each computational node is connected to at most $\alpha+1>2$ other nodes (an implementation constraint) with wires that allow for bidirectional communication. No other restriction is assumed on the topology of the decoder. In each iteration, each node sends messages to all its neighboring nodes. The maximum of all the neighborhood sizes (over all the message nodes) at the decoder of user i at the end of l_i iterations is denoted by $N_i \lesssim \alpha^{l_i}$. Each computational node is assumed to consume a fixed E_{node} joules of energy at each iteration. We define the parameter $\gamma = \frac{E_{node} h_1^2}{\sigma_0^2 \log_2(\alpha)}$ that captures the energy and the architecture terms relevant to our energy calculations.

B. Joint broadcast and time-division multiplexing strategies

We call a strategy a *joint broadcast* strategy if it requires each user to use the entire block to decode its own message. Superposition coding and dirty-paper coding might be considered as joint broadcast strategies. Also, we define a *time-division multiplexing* (TDM) strategy as one in which the signal for each user is sent at different time indices, and thus user i only uses indices assigned to itself to perform the decoding. We will analyze the performance of both of these strategies. The term *total energy* refers to the sum of the transmit energy E_T and the decoding energy $E_{dec}^{(i)}$. Our objective is to minimize the total energy per-bit, that is given by

$$E_{per-bit} = \frac{E_T + E_{dec}^{(1)} + E_{dec}^{(2)}}{k_1 + k_2}. \tag{2}$$

III. Lower bounds on the total energy for broadcast channels

A. Lower bounds on the total energy for joint broadcast schemes

The main result of this section is a lower bound on the total energy for joint broadcast strategies defined in Section II-B. It is presented in a sequence of three theorems. Theorem 1 derives lower bounds on the bit-error probabilities for the two users under some test channels. Theorem 2 uses results of Theorem 1 to derive lower bounds on neighborhood sizes (as in [7]) for given target error probabilities. These lower bounds are eventually used to derive to lower bounds on the total energy in Theorem 3.

Theorem 1: For a test channel pair (G, J) for the two-user broadcast channel of Section II, the following lower bounds hold on the error probabilities for all coding schemes that operate with average transmit power P and for all $\sigma_G^2 \leq \zeta\sigma_J^2$,

$$h_b(\langle P_e^{(1)}\rangle_G) \geq 1 - \frac{1}{2R_1}\log_2\left(1 + \frac{h_1^2 P_u^{\psi}}{\sigma_G^2}\right) =: \delta_1(P_u, \sigma_G^2)$$

$$h_b(\langle P_e^{(2)}\rangle_J) \geq 1 - \frac{1}{2R_2}\log_2\left(1 + \frac{h_2^2(P - P_u^{\psi})}{h_2^2 P_u^{\psi} + \sigma_J^2}\right) =: \delta_2(P_u, \sigma_J^2),$$

for some $0 \leq P_u^{\psi} \leq P$. Further, P_u^{ψ} is dependent only on the encoding strategy ψ, and not on the channels G and J.

Proof: We first prove the following two lemmas. The first lemma is a version of Fano's inequality [14, Pg. 79] for bit error probability. We include it for completeness.

Lemma 1: For a message $\mathbf{B}^k$ comprising of k bits communicated across a channel with rate R and output $\mathbf{Y}^m$, the average bit error probability $\langle P_e\rangle$ is lower bounded as follows

$$h_b(\langle P_e\rangle) \geq 1 - \frac{I(B^k;\mathbf{Y}^m)}{mR}. \tag{3}$$

Proof: Denoting by $\widehat{B}_i$ as the estimate of the i-th bit and $\widetilde{B}_i = \widehat{B}_i \oplus B_i$ as the binary error indicator,

$$\begin{aligned}
\frac{1}{k}H(\mathbf{B}^k|\mathbf{Y}^m) &= \frac{1}{k}H(\mathbf{B}^k \oplus \widehat{\mathbf{B}}^k|\mathbf{Y}^m) \\
&= \frac{1}{k}H(\widetilde{\mathbf{B}}^k|\mathbf{Y}^m) \\
&\leq \frac{1}{k}H(\widetilde{\mathbf{B}}^k) \\
&\leq \frac{1}{k}\sum_{i=1}^{k} H(\widetilde{B}_i) \\
&\overset{(a)}{=} \frac{1}{k}\sum_{i=1}^{k} h_b(p_i) \\
&\overset{(b)}{\leq} h_b(\langle P_e\rangle)
\end{aligned}$$

where in (a), $p_i = \Pr(\widetilde{B}_i = 1)$, and (b) follows from the concavity of $h_b(\cdot)$. ■

Lemma 2: Consider any two AWGN test channels G and J of variance σ_G^2 and σ_J^2 for user 1 and 2 respectively such that $\sigma_G^2 \leq \zeta\sigma_J^2$. For any encoding strategy with average channel input power P, the mutual information pair

$$\left(\frac{1}{m}I(\mathbf{B^{(1)}}^{k_1};\mathbf{Y}^m), \frac{1}{m}I(\mathbf{B^{(2)}}^{k_1};\mathbf{Z}^m)\right)$$

satisfies the following property.

$$\begin{aligned} \frac{1}{m}I(\mathbf{B^{(1)}}^{k_1};\mathbf{Y}^m) &\leq \frac{1}{2}\log_2\left(1+\frac{h_1^2P_u}{\sigma_G^2}\right) \\ \frac{1}{m}I(\mathbf{B^{(2)}}^{k_2};\mathbf{Z}^m) &\leq \frac{1}{2}\log_2\left(1+\frac{h_2^2(P-P_u)}{h_2^2P_u+\sigma_J^2}\right), \end{aligned} \tag{4}$$

for some $0 \leq P_u \leq P$.

Proof: Let $U_i = \left(\mathbf{Y}^{i-1}, \mathbf{B^{(2)}}^{k_2}\right)$, and let Q be uniformly distributed on $\{1,2,\ldots,m\}$. Define X,U,Y,Z by $X = X_i, U = U_i, Y = Y_i, Z = Z_i$ if $Q = i$. Following algebraic manipulations are based on those in [1, Pg. 455],

$$\begin{aligned} I(\mathbf{B^{(2)}}^{k_2};\mathbf{Z}^m) &= \sum_{i=1}^m I(\mathbf{B^{(2)}}^{k_2};Z_i|\mathbf{Z}^{i-1}) \\ &= \sum_{i=1}^m \left(h(Z_i|\mathbf{Z}^{i-1}) - h(Z_i|\mathbf{Z}^{i-1},\mathbf{B^{(2)}}^{k_2})\right) \\ &\stackrel{(a)}{=} \sum_{i=1}^m h(Z_i|\mathbf{Z}^{i-1}) - h(Z_i|\mathbf{Z}^{i-1},\mathbf{Y}^{i-1},\mathbf{B^{(2)}}^{k_2}) \\ &\stackrel{(b)}{\leq} \sum_{i=1}^m h(Z_i) - h(Z_i|\mathbf{Z}^{i-1},\mathbf{Y}^{i-1},\mathbf{B^{(2)}}^{k_2}) \\ &\stackrel{(c)}{\leq} \sum_{i=1}^m h(Z_i) - h(Z_i|\mathbf{Y}^{i-1},\mathbf{B^{(2)}}^{k_2}) \\ &\stackrel{(d)}{\leq} \sum_{i=1}^m I(U_i;Z_i) \\ &\stackrel{(e)}{=} mI(U;Z|Q), \end{aligned}$$

where (a) and (b) follow from the fact that conditioning reduces entropy, (c) follows from assuming that Z is a degraded version of Y, and hence there is a Markov chain $\mathbf{Z}^{i-1} - (\mathbf{B^{(2)}}^{k_2}, \mathbf{Y}^{i-1}) - Z_i$, (d) follows from definition of U_i and (e) from the definition of Q,U,Z. The assumption of degradedness of Z relative to Y is valid because of the condition that $\sigma_G^2 \leq \zeta\sigma_J^2$ and hence $\frac{h_1^2}{\sigma_G^2} > \frac{h_2^2}{\sigma_J^2}$.

Similarly, for user 1,

$$\begin{aligned} I(\mathbf{B^{(2)}}^{k_1};\mathbf{Y}^m) &\leq I(\mathbf{B^{(2)}}^{k_1};\mathbf{Y}^m,\mathbf{B^{(2)}}^{k_2}) \\ &\leq I(\mathbf{B^{(2)}}^{k_1};\mathbf{Y}^m|\mathbf{B^{(2)}}^{k_2}) \\ &= \sum_{i=1}^m I(\mathbf{B^{(2)}}^{k_1};Y_i|\mathbf{Y}^{i-1},\mathbf{B^{(2)}}^{k_2}) \\ &\leq \sum_{i=1}^m I(X_i;Y_i|U_i) \\ &= mI(X;Y|U,Q). \end{aligned}$$

Define the 'down-set' $\Delta\{(a,b)\}$ as follows

$$\Delta\{(a,b)\} := \{(x,y) : 0 \leq x \leq a, 0 \leq y \leq b\}. \tag{5}$$

Thus, the pair $\left(\frac{1}{m}I(\mathbf{B^{(1)}}^{k_1};\mathbf{Y}^m), \frac{1}{m}I(\mathbf{B^{(2)}}^{k_2};\mathbf{Y}^m)\right)$ is contained in $\bigcup_{(QU)-X-(YZ)}\Delta\{(I(X;Y|U,Q), I(U;Z|Q))\}$. For AWGN broadcast channel, it is well known (see, for example, [1, Pg. 427–428]) that this union is achieved by a jointly

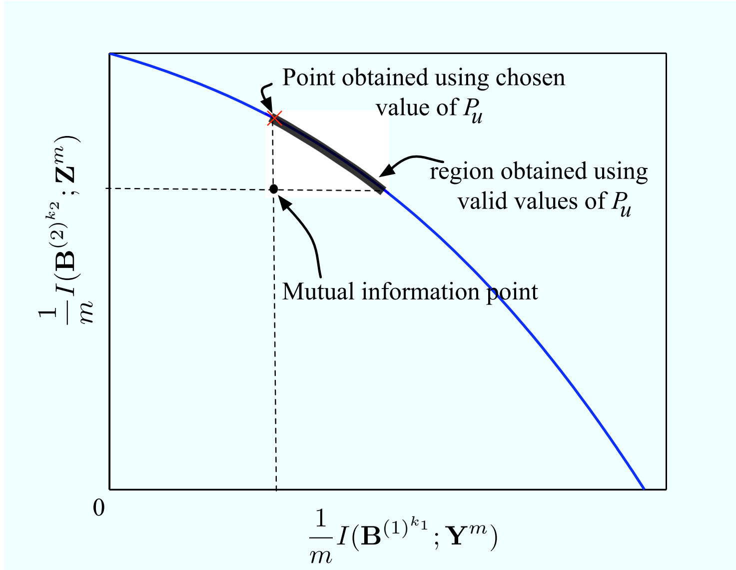

Fig. 1. From the converse argument, any achievable mutual information pair has a P_u that satisfies. The thick black curve is generated by the collection of P_u values each of which provides a valid upper bound to the mutual information pair. The point corresponding to the chosen value of P_u is shown by $\times$.

Gaussian choice of U and X, and Q constant. More precisely,

$$\bigcup_{(QU)-X-(YZ)} \Delta\{(I(X;Y|U,Q), I(U;Z|Q))\} = \bigcup_{\substack{U\sim\mathcal{N}(0,P_u)\\ V\sim\mathcal{N}(0,P-P_u),\text{ independent of } U\\ X=U+V, 0\le P_u\le P}} \Delta\{(I(X;Y|U,Q), I(U;Z|Q))\}$$
$$= \bigcup_{0\le P_u\le P} \Delta\left\{\left(\frac{1}{2}\log_2\left(1+\frac{h_1^2P_u}{\sigma_G^2}\right), \frac{1}{2}\log_2\left(1+\frac{h_2^2(P-P_u)}{h_2^2P_u+\sigma_J^2}\right)\right)\right\}. \quad (6)$$

Thus there exists a P_u such that the mutual information pair

$$\left(\frac{1}{m}I(\mathbf{B}^{(\mathbf{1})^{k_1}};\mathbf{Y}^m), \frac{1}{m}I(\mathbf{B}^{(\mathbf{2})^{k_1}};\mathbf{Z}^m)\right) \in \Delta\left\{\left(\frac{1}{2}\log_2\left(1+\frac{h_1^2P_u}{\sigma_G^2}\right), \frac{1}{2}\log_2\left(1+\frac{h_2^2(P-P_u)}{h_2^2P_u+\sigma_J^2}\right)\right)\right\}. \quad (7)$$

This completes the proof of Lemma 2. ∎

Continuing with the proof of Theorem 1, for any choice of the encoding function $\psi(\cdot)$, σ_G^2 and σ_J^2, there may exist more than one value of P_u such that (7) is satisfied, as shown in Fig. 1. Choose the value of P_u that satisfies $\frac{1}{m}I(\mathbf{B}^{(\mathbf{1})^{k_1}};\mathbf{Y}^m) = \frac{1}{2}\log_2\left(1+\frac{h_1^2P_u}{\sigma_G^2}\right)$. Denote this value of P_u by $P_u^\psi(\sigma_G^2,\sigma_J^2)$. Now the function $P_u^\psi(\sigma_G^2,\sigma_J^2)$ is well-defined for the given encoding function ψ as long as $\sigma_G^2 \le \zeta\sigma_J^2$. From Lemma 1 and 2, it follows that

$$h_b(\langle P_e^{(1)}\rangle_G) \ge 1-\frac{1}{2R_1}\log_2\left(1+\frac{h_1^2P_u^\psi(\sigma_G^2,\sigma_J^2)}{\sigma_G^2}\right) \quad (8a)$$

$$h_b(\langle P_e^{(2)}\rangle_J) \ge 1-\frac{1}{2R_2}\log_2\left(1+\frac{h_2^2(P-P_u^\psi(\sigma_G^2,\sigma_J^2))}{h_2^2P_u^\psi(\sigma_G^2,\sigma_J^2)+\sigma_J^2}\right). \quad (8b)$$

Thus the upper bounds on Theorem 1 hold for $P_u^\psi(\sigma_G^2,\sigma_J^2)$ that can in general be a function of σ_G^2 and σ_J^2. The following argument, which is key to obtaining a good bound, shows that the upper bounds hold for some P_u^ψ that is dependent only on the coding scheme.

First observe that (8a) can be loosened to

$$h_b(\langle P_e^{(1)}\rangle_G) \ge 1-\frac{1}{2R_1}\log_2\left(1+\frac{h_1^2\sup_{\tilde\sigma_G^2\le\zeta\sigma_J^2}P_u^\psi(\tilde\sigma_G^2,\sigma_J^2)}{\sigma_G^2}\right). \quad (9)$$

Since the noises at the two receivers are independent, and the scheme is designed with the knowledge of only σ_0^2, $\langle P_e^{(1)} \rangle_G$ does not depend on the channel J to user 2. Thus,

$$h_b(\langle P_e^{(1)} \rangle_G) = \sup_{\widetilde{\sigma}_J^2} h_b(\langle P_e^{(1)} \rangle_G) \geq 1 - \frac{1}{2R_1} \log_2 \left(1 + \frac{h_1^2 \inf_{\widetilde{\sigma}_J^2} \sup_{\widetilde{\sigma}_G^2 \leq \zeta \widetilde{\sigma}_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \widetilde{\sigma}_J^2)}{\sigma_G^2} \right). \tag{10}$$

Now consider (8b). Since $\langle P_e^{(2)} \rangle_J$ does not depend on the channel G to user 1,

$$\begin{aligned} h_b(\langle P_e^{(2)} \rangle_J) &= \sup_{\widetilde{\sigma}_G^2 \leq \zeta \sigma_J^2} h_b(\langle P_e^{(2)} \rangle_J) \\ &\overset{\text{(using (8b))}}{\geq} \sup_{\widetilde{\sigma}_G^2 \leq \zeta \sigma_J^2} \left(1 - \frac{1}{2R_2} \log_2 \left(1 + \frac{h_2^2 (P - P_u^{\psi}(\widetilde{\sigma}_G^2, \sigma_J^2))}{h_2^2 P_u^{\psi}(\widetilde{\sigma}_G^2, \sigma_J^2) + \sigma_J^2} \right) \right) \\ &= 1 - \frac{1}{2R_2} \log_2 \left(1 + \frac{h_2^2 (P - \sup_{\widetilde{\sigma}_G^2 \leq \zeta \sigma_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \sigma_J^2))}{h_2^2 \sup_{\widetilde{\sigma}_G^2 \leq \zeta \sigma_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \sigma_J^2) + \sigma_J^2} \right). \end{aligned}$$

Again, we loosen the bound

$$h_b(\langle P_e^{(2)} \rangle_J) \geq 1 - \frac{1}{2R_2} \log_2 \left(1 + \frac{h_2^2 (P - \inf_{\widetilde{\sigma}_J^2} \sup_{\widetilde{\sigma}_G^2 \leq \zeta \widetilde{\sigma}_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \widetilde{\sigma}_J^2))}{h_2^2 \inf_{\widetilde{\sigma}_J^2} \sup_{\widetilde{\sigma}_G^2 \leq \zeta \widetilde{\sigma}_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \widetilde{\sigma}_J^2) + \sigma_J^2} \right). \tag{11}$$

From (10) and (11), it follows that there exists a $P_u^{\psi} := \inf_{\widetilde{\sigma}_J^2} \sup_{\widetilde{\sigma}_G^2 \leq \zeta \widetilde{\sigma}_J^2} P_u^{\psi}(\widetilde{\sigma}_G^2, \widetilde{\sigma}_J^2)$ that is independent of σ_G^2, σ_J^2 and it satisfies (8). This completes the proof of Theorem 1.

■

Theorem 2 derives lower bounds on the error probabilities $\langle P_e^{(i)} \rangle_0$ for the underlying channel as a function of the neighborhood sizes at the two decoders. Turned around, these bounds provide lower bounds on the required neighborhood size N_i for given error probabilities. Using $l_i \geq \frac{\log_2(N_i)}{\log_2(\alpha)}$, the theorem provides lower bounds on the number of iterations. The derivation uses the following lemma from [7, Lemma 10].

Lemma 3: Let the underlying AWGN channel be of noise variance σ_0^2. Consider a test channel G of noise variance $\sigma_G^2 > \sigma_0^2$. Let A be a set of noise realizations $\mathbf{w}^n$ of length n such that $\Pr_G(\mathbf{w}^n \in A) = \delta$. Then,

$$\Pr_0(\mathbf{w}^n \in A) \geq f_G(n, \delta), \tag{12}$$

where,

$$\begin{aligned} f_G(n, x) : &= \frac{x}{2} \exp \Bigg(-nD(\sigma_G^2 || \sigma_0^2) \\ &\quad -\sqrt{n} \left(\frac{3}{2} + 2 \ln \left(\frac{2}{x} \right) \right) \left(\frac{\sigma_G^2}{\sigma_0^2} - 1 \right) \Bigg). \end{aligned}$$

Further, $f_G(n, \cdot)$ is a convex-$\cup$ increasing function for any fixed n and for all values of $\sigma_G^2 \geq \sigma_0^2$.

Proof: See [7, Lemma 10]. ■

Theorem 2: For AWGN broadcast channel with total average input power P, the following pair of equations provide lower bounds on the neighborhood sizes at the two decoders for the decoding model of Section II-A for given bit-error probabilities $\langle P_e^{(i)} \rangle_0$ at the two users.

$$\begin{aligned} \langle P_e^{(1)} \rangle_0 &\geq f_G \left(N_1, \frac{h_b^{-1}(\delta_1(P_u, \sigma_G^2))}{2} \right) \\ \langle P_e^{(2)} \rangle_0 &\geq f_J \left(N_2, \frac{h_b^{-1}(\delta_2(P_u, \sigma_J^2))}{2} \right) \end{aligned}$$

for all σ_G^2, σ_J^2 satisfying $\sigma_G^2 < \zeta \sigma_J^2$ and for some constant $P_u \in [0, P]$ that depends only on the coding scheme. Here $\delta_i(\sigma^2)$ are as defined in Theorem 1.

Proof: Using Theorem 1, under test channels G and J for user 1 and 2 respectively,

$$\begin{aligned} h_b(\langle P_e^{(1)} \rangle_G) &\geq 1 - \frac{1}{2R_1} \log_2 \left(1 + \frac{h_1^2 P_u}{\sigma_G^2} \right) = \delta_1(P_u, \sigma_G^2) \\ h_b(\langle P_e^{(2)} \rangle_J) &\geq 1 - \frac{1}{2R_2} \log_2 \left(1 + \frac{h_2^2 (P - P_u)}{\sigma_J^2} \right) = \delta_2(P_u, \sigma_J^2), \end{aligned} \tag{13}$$

for some P_u. Now consider the i-th message bit $B_i^{(1)}$. The decoding at receiver 1 is performed based on observations in a decoding neighborhood $\mathbf{y}_{nbd,i}^{N_1}$. An error happens if the noise at the first receiver, $\mathbf{w}_{nbd,i}^{N_1}$, lies in a region $\mathcal{D}(\mathbf{x}_{nbd,i}^{N_1}, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2})$ that pushes $\mathbf{y}_{nbd,i}^{N_1}$ into a region such that the decoder makes an error in decoding $B_i^{(1)}$. Thus

$$\langle P_e^{(1)} \rangle_G = \frac{1}{k_1} \sum_i \frac{1}{2^{k_1+k_2}} \sum_{{b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}} \text{Pr}_G(\mathbf{w}_{nbd,i}^{N_1} \in \mathcal{D}(\mathbf{x}_{nbd,i}^{N_1}, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2})). \tag{14}$$

Now the average error probability for the underlying channel of noise variance σ_0^2 is given by

$$\begin{aligned}
\langle P_e^{(1)} \rangle_0 &= \frac{1}{k_1 2^{k_1+k_2}} \sum_{i, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}} \text{Pr}_0(\mathbf{w}_{nbd,i}^{N_1} \in \mathcal{D}(\mathbf{x}_{nbd,i}^{N_1}, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2})) \\
&\overset{(a)}{\geq} \frac{1}{k_1 2^{k_1+k_2}} \sum_{i, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}} f_G\left(N_1, \text{Pr}_G(\mathbf{w}_{nbd,i}^{N_1} \in \mathcal{D}(\mathbf{x}_{nbd,i}^{N_1}, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}))\right) \\
&\overset{(b)}{\geq} f_G\left(N_1, \frac{1}{k_1 2^{k_1+k_2}} \sum_{i, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}} \text{Pr}_G(\mathbf{w}_{nbd,i}^{N_1} \in \mathcal{D}(\mathbf{x}_{nbd,i}^{N_1}, {b^{(1)}}^{k_1}, {b^{(2)}}^{k_2}))\right) \\
&\overset{(c)}{\geq} f_G(N_1, \langle P_e^{(1)} \rangle_G),
\end{aligned}$$

where (a) is obtained using Lemma 3, (b) follows from convexity of $f_G(N_1, \cdot)$ (Lemma 3), and (c) follows from (14) and the monotonicity of $f_G(N_1, \cdot)$. The proof now follows from (13). The lower limit on σ_G^2 is obtained from Theorem 1.

The proof of the lower bound on $\langle P_e^{(2)} \rangle$ in Theorem 2 follows *mutatis mutandis* and is hence omitted.

■

Finally, Theorem 3 derives lower bounds on the total energy per-bit for joint broadcast strategies using the lower bounds on the number of iterations obtained from Theorem 2.

Theorem 3 (Energy per-bit for joint strategies): The total per-bit energy required by any joint broadcast strategy for communicating at rates R_1, R_2 to the two users is lower bounded as follows.

$$\begin{aligned}
\frac{E_{per-bit} h_1^2}{\sigma_0^2} \geq \min_P \Bigg\{ & \frac{P h_1^2}{(R_1+R_2)\sigma_0^2} + \gamma \min_{P_u} \Bigg(\max_{\sigma_G^2} \Bigg\{ \\
& \frac{R_1+1}{R_1+R_2} \log_2\left(N_1(P, \sigma_G^2, P_u)\right) + \frac{R_2+1}{R_1+R_2} \times \\
& \max_{\sigma_J^2 > \frac{\sigma_G^2}{\zeta}} \left\{ \log_2\left(N_2(P, \sigma_J^2, P_u)\right) \right\} \Bigg\} \Bigg) \Bigg\},
\end{aligned}$$

where the functions N_1 and N_2 are lower bounded as in Theorem 2, and the optimization is over

$$\begin{aligned}
P &\geq \frac{1}{h_1^2}(2^{2R_1(1-h_b(\langle P_e^{(1)} \rangle_0))} - 1) 2^{2R_2(1-h_b(\langle P_e^{(2)} \rangle_0))} \sigma_0^2 \\
&\quad + \frac{1}{h_2^2}(2^{2R_2(1-h_b(\langle P_e^{(2)} \rangle_0))} - 1)\sigma_0^2,
\end{aligned}$$

and P_u satisfying $h_b(\langle P_e^{(i)} \rangle_0) \geq \delta_i(\sigma_0^2)$ for $i = 1, 2$.

Proof: The number of nodes at decoder i is lower bounded by $k_i + m$, the sum of the number of output nodes and number of message nodes. The energy consumed at decoder i is therefore lower bounded by $E_{node}(k_i + m) l_i$, where l_i is the number of iterations at decoder i.

$$\begin{aligned}
E_{tot} &\geq \min_P mP + E_{node}((k_1+m)l_1 + (k_2+m)l_2) \\
\Rightarrow \frac{E_{tot} h_1^2}{\sigma_0^2 (k_1+k_2)} &\geq \min_P \frac{m}{k_1+k_2} \frac{P h_1^2}{\sigma_0^2} + \gamma \min_{P_u} \max_{\sigma_G^2, \sigma_J^2 : \sigma_G^2 \leq \zeta \sigma_J^2} \Bigg(\frac{k_1+m}{k_1+k_2} \log_2\left(N_1(P, \sigma_G^2, P_u)\right) \\
&\quad + \frac{k_2+m}{k_1+k_2} \log_2\left(N_2(P, \sigma_J^2, P_u)\right) \Bigg)
\end{aligned}$$

The lower bound on P is the minimum transmit power required to communicate at rates R_i with error probabilities $\langle P_e^{(i)} \rangle$ for the underlying channels. It is obtained from Theorem 1, using G and J as the underlying channel. The limits on P_u are obtained from Lemma 1, again using G and J as the underlying channel.

■

B. Lower bounds on the total energy for time-division multiplexing

Theorem 4 (Energy per-bit for TDM): The total energy per-bit for the time-division multiplexing scheme (under the model described in Section II-A) that communicates k_i bits to user i at rate R_i (so that $\frac{k_1}{R_1} = \frac{k_2}{R_2}$) is lower bounded by

$$\begin{aligned} \frac{(R_1+R_2)E_{per-bit}h_1^2}{\sigma_0^2} &\geq \min_{\widetilde{R}_1, P_1, P_2} \left\{ \frac{P_1 h_1^2 R_1}{\sigma_0^2 \widetilde{R}_1} + \frac{P_2 h_1^2 R_2}{\sigma_0^2 \widetilde{R}_2} \right. \\ + &\left. \gamma \frac{(\widetilde{R}_1+1)R_1}{\widetilde{R}_1} \log_2(N_1) + \gamma \frac{(\widetilde{R}_2+1)R_2}{\widetilde{R}_2} \log_2(N_2) \right\}, \end{aligned}$$

where $\widetilde{R}_1 \geq R_1$, and $\widetilde{R}_2$ is given by

$$\frac{R_1}{\widetilde{R}_1} + \frac{R_2}{\widetilde{R}_2} = 1, \tag{15}$$

$h_i^2 P_i \geq (2^{2\widetilde{R}_i(1-h_b(\langle P_e^{(i)}\rangle_0))} - 1)\sigma_0^2$, and N_i is lower bounded as follows

$$\langle P_e^{(i)} \rangle_0 \geq f_i \left(N_i, \frac{h_b^{-1}(\delta^{(i)}(\sigma_i^2))}{2} \right),$$

where

$$\delta^{(i)}(\sigma_i^2) = 1 - \frac{1}{2\widetilde{R}_i} \log_2 \left(1 + \frac{h_i^2 P_i}{\sigma_i^2} \right),$$

for all σ_i^2 satisfying $\delta^{(i)}(\sigma_i^2) > 0$. These lower bounds can be optimized over σ_i^2.

Proof: $\widetilde{R}_1$ and $\widetilde{R}_2$ are the codebook rates of the two users. However the effective rates are lower because the total channel uses m required is the sum of the channel uses m_i that the two user require. This gives us the following relation

$$\begin{aligned} R_2 &= \frac{k_2}{m_1+m_2} \\ \Rightarrow \frac{1}{R_2} &= \frac{m_1}{k_2} + \frac{m_2}{k_2} \\ &= \frac{m_1}{k_2} + \frac{1}{\widetilde{R}_2}. \end{aligned}$$

Thus,

$$\frac{m_1}{k_2} = \frac{1}{R_2} - \frac{1}{\widetilde{R}_2}. \tag{16}$$

Similarly, $\frac{m_2}{k_1} = \frac{1}{R_1} - \frac{1}{\widetilde{R}_1}$. Also,

$$\begin{aligned} \frac{m_1}{k_2} &= \frac{k_1}{\widetilde{R}_1 R_2 m} \\ &= \frac{R_1}{\widetilde{R}_1 R_2}. \end{aligned} \tag{17}$$

From (16) and (17),

$$\frac{R_1}{\widetilde{R}_1} + \frac{R_2}{\widetilde{R}_2} = 1. \tag{18}$$

The total energy is the sum of the transmit energy and the decoding energy. The total transmit energy is given by $m_1 P_1 + m_2 P_2$. The number of nodes at user i is lower bounded by $k_i + m_i$. Thus, the total energy is lower bounded as follows

$$\begin{aligned} E_{tot} &\geq m_1P_1 + m_2P_2 + (k_1+m_1)E_{node}l_1 + (k_2+m_2)E_{node}l_2 \\ \frac{E_{tot}}{k_1+k_2} &\geq \frac{m_1P_1}{k_1+k_2} + \frac{m_2P_2}{k_1+k_2} + \frac{k_1+m_1}{k_1+k_2}E_{node}l_1 + \frac{k_2+m_2}{k_1+k_2}E_{node}l_2 \\ &= \frac{P_1}{\frac{k_1}{m_1}+\frac{k_2}{m_1}} + \frac{P_2}{\frac{k_1}{m_2}+\frac{k_2}{m_2}} + \frac{\frac{k_1}{m_1}+1}{\frac{k_1}{m_1}+\frac{k_2}{m_1}}E_{node}l_1 + \frac{\frac{k_2}{m_2}+1}{\frac{k_1}{m_2}+\frac{k_2}{m_2}}E_{node}l_2 \\ &= \frac{P_1R_1}{\widetilde{R}_1(R_1+R_2)} + \frac{P_2R_2}{\widetilde{R}_2(R_1+R_2)} + \frac{(\widetilde{R}_1+1)R_1}{\widetilde{R}_1(R_1+R_2)}E_{node}l_1 + \frac{(\widetilde{R}_2+1)R_2}{\widetilde{R}_2(R_1+R_2)}E_{node}l_2. \end{aligned}$$

The result now follows from normalization.

■

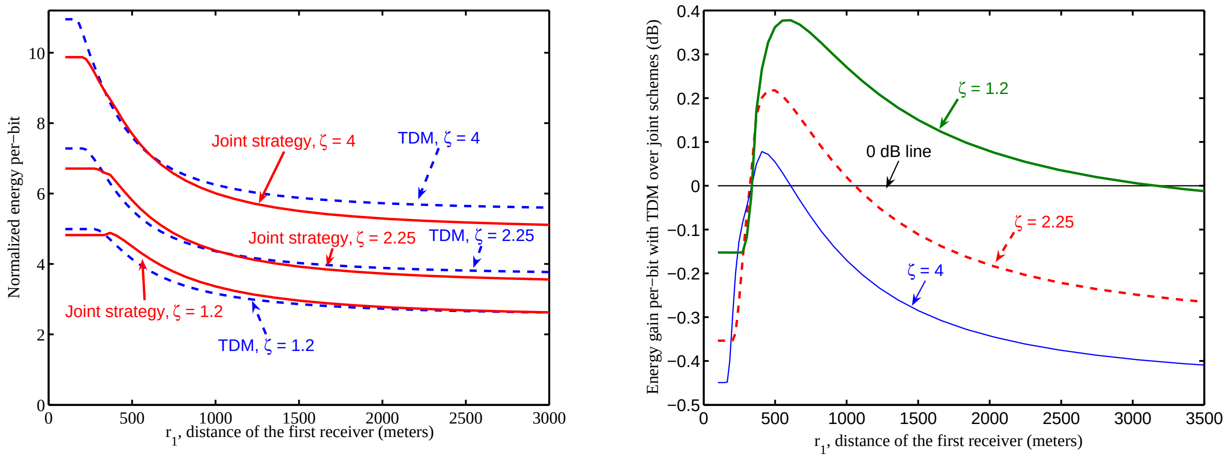

Fig. 2. The first plot shows the lower bounds on the (normalized) energy per-bit vs distance of the first user for joint broadcast schemes and TDM for different values of $\zeta = r_2^2/r_1^2$ for $\langle P_e^{(i)} \rangle_0 = 10^{-9}$. The total energy per-bit can be smaller for TDM for short distances and reasonably small values of ζ. The second plot shows the difference in energy per-bit (in dB) for TDM vs joint broadcast schemes. At moderate distances, the gap can be large for ζ close to 1, while the advantage of TDM is small at large ζ. The plots are for a transmit frequency of 3 GHz, an E_{node} value of 1 pJ, $\sigma_0^2 = \kappa T$ with κ as Boltzmann's constant and $T = 300$ K, and the y-axis shows the normalized energy $\frac{E_{per-bit} h_1^2}{\sigma_0^2}$. The figure assumes $k_1 = k_2$, and $R_1 = R_2 = 1/3$.

C. Performance comparison

In this and the following section we assume $R_i = R_{des} = 1/3$ (the desired rate), $k_1 = k_2$, $E_{node} = 1$ pJ, temperature $T = 300$ K, and an operating frequency is 3 GHz. Fig. 2 shows a comparison of the normalized total energy per-bit (given by $\frac{E_{per-bit} h_1^2}{\sigma_0^2}$) for various values of $r_1$and $\zeta = \frac{r_2^2}{r_1^2}$. For small ζ, the performance gain of TDM is substantially better than any joint scheme for distances as large as 3000 m for the given system parameters. To understand why this must be the case, consider $\zeta = 1$, the case of equal fade coefficients. It turns out that in this case, TDM also achieves the capacity of the Gaussian broadcast channel, and is hence transmit-power optimal. Intuitively, since TDM makes more efficient use of decoding energy, it should require smaller total energy for $\zeta = 1$.

For larger ζ, the required transmit power using joint schemes vis-a-vis that required by TDM is much smaller, and joint schemes start dominating TDM (see Fig. 3). This is particularly true at large distances, or high error probabilities, where decoding energy ceases to matter. Again, if error probability is lowered for fixed ζ (see Fig. 4), TDM outperforms joint broadcast schemes because TDM's savings in decoding energy exceed its spending on transmit energy.

Even though Fig. 2 suggests that joint schemes dominate TDM at small distances (~ 300 m or smaller at 3 GHz, depending on ζ), we believe that this is a consequence of the looseness in our bounds. An increase in transmit power can force the lower bounds on the neighborhood sizes to 1, thereby making the lower bound on the decoding energy zero, even though the actual decoding energy itself is non-zero.

D. Allowing flexibility in the communication rate

For point-to-point communication, [13] suggests that for a given target error probability, the optimal rate R^* that minimizes the total energy per-bit can be non-zero. Consequently, it is desirable to have the operating rate $R = R^*$ if $R^* > R_{des}$.

To see if this improves the performance of the joint broadcast schemes, we allow for flexibility in R subject to $R > R_{des}$. Fig. 5 shows that at very short distances, the total power is an increasing function of the rate, and hence $R^* = 0$ (this happens again due to the looseness of our bounds). At extremely large distances, the decoding energy is irrelevant, and R^* approaches 0. At moderate distances, for the system parameters in Fig. 2, 3 and 4, R^* for joint broadcast schemes turns out to be smaller than $R_{des} = 1/3$ (see Fig. 5), and hence the same plots are observed even allowing for the rate flexibility. Though there would be an impact of flexible rate at extremely low error probabilities, it would benefit both TDM and joint broadcast schemes, and thus needs further study.

IV. Improved sphere-packing bounds for the Gaussian broadcast problem

Outer bounds on the error exponents have appeared in [15]. The bounds are obtained by considering the point-to-point channels to the two users, and point-to-point channel obtained by allowing the two users to cooperate. This bound is shown (as the dashed-curve) in Fig. 6.

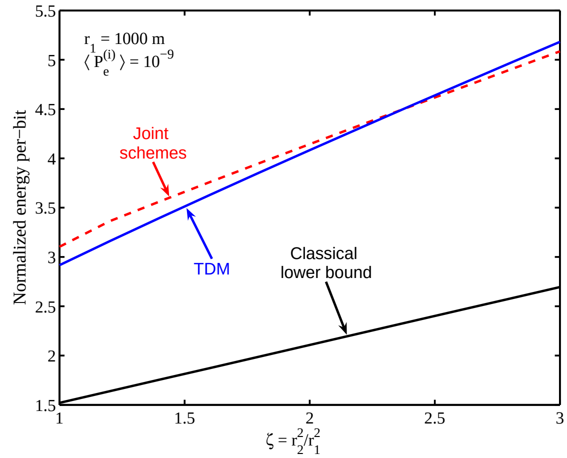

Fig. 3. The plot shows the lower bounds on the (normalized) energy per-bit vs $\langle P_e^{(i)} \rangle_0$ for joint strategies and TDM for $\langle P_e^{(i)} \rangle_0 = 10^{-9}$. For ζ close to 1, TDM requires smaller energy than the joint broadcast schemes. For ζ much larger than 1, the joint broadcast schemes outperform TDM due to improved savings in transmit energy. The other parameters are the same as in Fig. 2. The classical lower bound disregards the decoding energy.

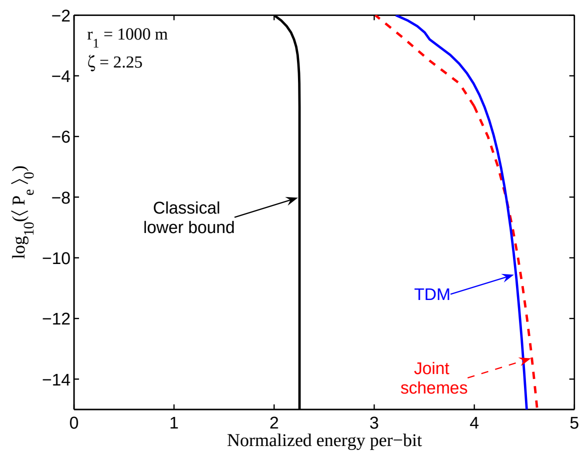

Fig. 4. The plot shows the lower bounds on the (normalized) energy per-bit vs $\langle P_e^{(i)} \rangle_0$ (equal at the two receivers) for joint strategies and TDM for $\zeta = r_2^2/r_1^2 = 2.25$ at $r_1 = 1000$ m. As the error probability decreases, the required decoding energy increases for either scheme. Since TDM saves on decoding energy, it outperforms joint broadcast schemes at low $\langle P_e^{(i)} \rangle_0$. The other parameters are the same as that in Fig. 2.

Theorem 2 can be used to derive outer bounds as follows. Since P_u^{ψ} does not depend on (σ_G^2, σ_J^2), the outer bounds can be expressed as follows.

$$(E_1, E_2) \in \bigcup_{P_u^{\psi} \geq 0} \bigcap_{\substack{\zeta \sigma_J^2 \geq \sigma_G^2 \geq \sigma_0^2 : \\ \frac{1}{2}\log_2\left(1 + \frac{h_2^2 (P - P_u^{\psi})}{h_2^2 P_u^{\psi} + \sigma_J^2}\right) \leq R_1,\ \frac{1}{2}\log_2\left(1 + \frac{h_1^2 P_u^{\psi}}{\sigma_J^2}\right) \leq R_2}} \Delta\left\{\left(D(\sigma_G^2 || \sigma_0^2), D(\sigma_J^2 || \sigma_0^2)\right)\right\}. \tag{19}$$

The reasoning is the following: for any given P_u^{ψ}, we are free to choose σ_G^2 and σ_J^2 (as long as the degradation is not reversed) to obtain the best bound. Further, another bound is obtained by considering the channel with degradation reversed, i.e. $\zeta^2 \sigma_J^2 < \sigma_G^2$.

$$(E_1, E_2) \in \bigcup_{P_u^{\psi} \geq 0} \bigcap_{\substack{\sigma_G^2 \geq \zeta \sigma_J^2 \geq \sigma_0^2 : \\ \frac{1}{2}\log_2\left(1 + \frac{h_2^2 P_u^{\psi}}{\sigma_J^2}\right) \leq R_1,\ \frac{1}{2}\log_2\left(1 + \frac{h_1^2 (P - P_u^{\psi})}{h_1^2 P_u^{\psi} + \sigma_J^2}\right) \leq R_2}} \Delta\left\{\left(D(\sigma_G^2 || \sigma_0^2), D(\sigma_J^2 || \sigma_0^2)\right)\right\}. \tag{20}$$

Taking the intersection of the regions defined by (19) and (20), we obtain a tighter bound on the error exponent region, which is plotted in Fig. 6 for the parameter choice of [15]. It turns out that this improvement in bounds is necessary to obtain the

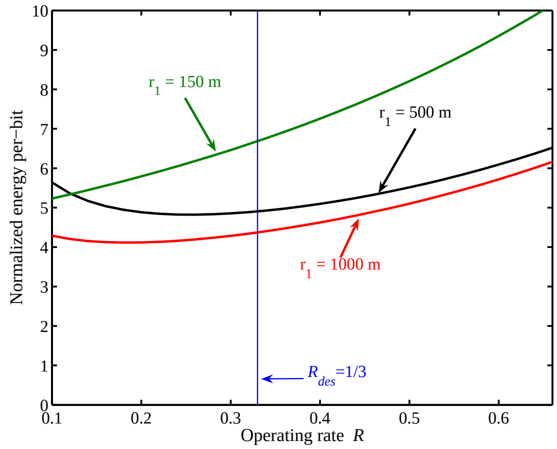

Fig. 5. For joint strategies at moderately large error probabilities (10^{-9} in this plot), the optimal communication rate R^* that minimizes our lower bounds is 0 at short distances, and close to 0 at large distances. At moderate distances, R^* can be large but is observed to be smaller than $R_{des} = 1/3$ for the cases considered in this paper.

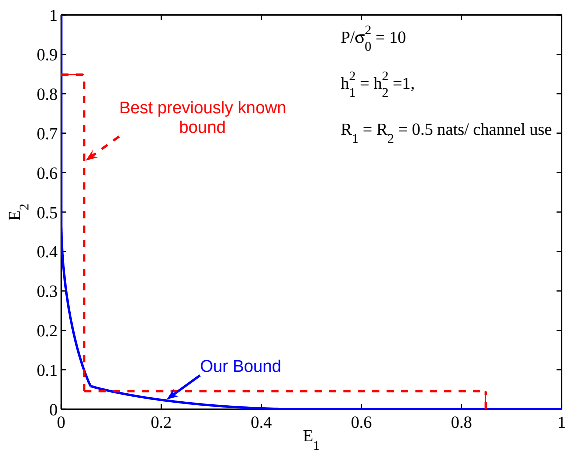

Fig. 6. Comparison of bounds on error exponents from proof of Theorem 2 and the best known bounds of [15] for the Gaussian broadcast problem.

strong results in energy consumption described in earlier parts of the paper.

V. Discussions and Conclusions

We note that the lower bounds on the joint broadcast schemes here are optimistic because they assume that user 1 can decode its own message without decoding any part of the message for user 2. In practice, for the two well known joint broadcast schemes of superposition and DPC, user 1 decodes more bits than merely its own message bits. In superposition, user 1 decodes the entire message for user 2. In DPC, user 2 decodes an auxiliary codeword. The number of such auxiliary codewords far exceeds the number of possible messages for user 1. Thus for either superposition or DPC, the decoding energy at the first user is higher than that assumed in Theorem 3 because the nodes dedicated to decoding these extra bits consume energy as well.

The minimum energy communication scheme for broadcast channels at short distances must have some aspect of TDM, that is, each receiver should not require the entire block for decoding. We believe that the optimal scheme would time-share between a joint scheme and TDM, thus balancing between loss in the rate in TDM and increase in the decoding power in the joint schemes.

Acknowledgments

We thank Amin Gohari for useful discussions. We gratefully acknowledge the support of the National Science Foundation (CNS-403427 and CCF-729122) and Sumitomo Electric.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 1st ed. New York: Wiley, 1991.

[2] A. Avestimehr, S. Diggavi, and D. Tse, “A deterministic approach to wireless relay networks,” in *Proc. of the Allerton Conference on Communications, Control and Computing*, October 2007.

[3] A. Avestimehr, “Wireless network information flow: A deterministic approach,” Ph.D. dissertation, UC Berkeley, Berkeley, CA, 2008.

[4] A. Lapidoth and S. M. Moser, “Capacity bounds via duality with applications to multiple-antenna systems on flat-fading channels,” *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 1019–1030, Oct. 2003.

[5] S Cui, AJ Goldsmith and A Bahai, “Energy Constrained Modulation Optimization,” *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 1–11, 2005.

[6] A. J. Goldsmith and S. B. Wicker, “Design challenges for energy constrained ad hoc wireless networks,” *IEEE Trans. Wireless Commun.*, pp. 8–27, 2002.

[7] A. Sahai and P. Grover, “The price of certainty : “waterslide curves” and the gap to capacity,” Dec. 2007. [Online]. Available: http://arXiv.org/abs/0801.0352v1

[8] P. Massaad, M. Medard, and L. Zheng, “Impact of Processing Energy on the Capacity of Wireless Channels,” in *International Symposium on Information Theory and its Applications (ISITA)*, 2004.

[9] S. Verdu, “On channel capacity per unit cost,” *IEEE Trans. Inform. Theory*, vol. 36, pp. 1019–1030, 1990.

[10] P. Massaad, M. Medard, and L. Zheng, “On the capacity of multiple-access channels with processing power,” in *Proceedings of the Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Oct. 2004.

[11] M. Lentmaier, D. V. Truhachev, K. S. Zigangirov, and D. J. Costello, “An analysis of the block error probability performance of iterative decoding,” *IEEE Trans. Inform. Theory*, vol. 51, no. 11, pp. 3834–3855, Nov. 2005.

[12] T. Richardson and R. Urbanke, *Modern Coding Theory*. Cambridge University Press, 2007.

[13] P. Grover and A. Sahai, “Green codes: Energy-efficient short-range communication,” in *Proceedings of the 2008 IEEE Symposium on Information Theory*, Toronto, Canada, Jul. 2008.

[14] R. G. Gallager, *Information Theory and Reliable Communication*. New York, NY: John Wiley, 1971.

[15] L. Weng, S. S. Pradhan, and A. Anastasopoulos, “Error exponent regions for gaussian broadcast and multiple access channels,” *IEEE Trans. Inform. Theory*, vol. 54, no. 7, pp. 2919–2942, Jul. 2008.