Cross-layer Design of Energy-constrained Networks Using Cooperative MIMO Techniques

Shuguang Cui and Andrea J. Goldsmith
Wireless Systems Lab,
Dept. of EE, Stanford University,
Stanford, CA 94305
Email: {shuguang, andrea}@systems.stanford.edu

Abstract

We consider networks where energy is a limited resource so that energy consumption must be minimized while satisfying given throughput requirements. For such networks, cross-layer design coupled with node cooperation can significantly reduce both energy consumption and delay. In this paper, we propose a cooperative multiple-input multiple-output (MIMO) technique where multiple nodes within a cluster cooperate in signal transmission and/or reception. In our scheme, local information exchange within the cluster is not necessary if Alamouti codes are used with appropriate transmission scheduling. A cross-layer design framework is then applied that optimizes routing to minimize the energy consumption and delay. For the cooperative MIMO scheme, routing is optimized based on an equivalent single-input single-out (SISO) system, where each cooperating cluster is treated as a super node. We derive the best energy-delay tradeoff curve under this optimization model and show that the cooperative MIMO approach dramatically improves energy and delay performance, especially when link layer adaptation is used.

Index Terms

Energy efficiency, joint routing and scheduling, link adaptation, cooperative MIMO, cross-layer.

This research is supported in part by funds from National Semiconductor, the Robert Bosch Corporation, and ONR grant N00014-05-1-0168.

I. INTRODUCTION

Energy-constrained networks, such as sensor networks, have nodes typically powered by small batteries, for which replacement or recharging is very difficult if not impossible. As a result, minimizing the energy per bit required for reliable end-to-end transmission becomes an important design consideration. For short range applications, transmission energy may no longer be dominant. Thus, transmission and circuit processing energy must be jointly minimized [1]. As such, we focus on minimizing the total energy consumption given certain network throughput requirements and delay constraints. Since all the layers in the protocol stack affect the total energy consumption, throughput, and delay, cross-layer design is necessary for energy minimization [2].

Cross-layer design for improving the network performance has been a focus of much recent work. Joint scheduling and power control to reduce energy consumption and increase single hop throughput is considered in [3]. Cross-layer design based on computation of optimal power control, link schedule, and routing flow is described in [4]. The aim of that paper is to minimize the average transmission power over an infinite horizon. Also, the routing flow is computed in an incremental manner: it uses the Lagrange multipliers obtained at each step by solving an optimization problem of possibly exponential complexity in the number of links. Energy efficient power control and scheduling, with no rate adaptation on links, is considered in [5]. Joint routing, power control, and scheduling for a TDMA-CDMA network is investigated in [6] and [7]. However, in all of this work, only transmission energy is considered and hardware processing energy is ignored. This can lead to suboptimal performance in short range networks.

Energy minimization including hardware constraints is investigated in [8]-[10], where the authors proposes a joint design between the link layer and the silicon layer. By considering constraints such as power consumption imposed by the underlying circuits, optimal modulation schemes are derived to minimize the total energy consumption. However, these results do not take into account higher layer protocols such as MAC and routing. In [11], joint routing, MAC, and link layer optimization techniques to minimize the sum of the transmission energy and the circuit processing energy are proposed, where the flow control and link scheduling problem is approximated as a convex problem and efficient optimization methods are applied to find the solution.

In the physical layer, multiple antenna techniques have been shown to be very effective in improving the performance of wireless systems in the presence of fading [12], where the performance gain is in the form of diversity gain, array gain, and multiplexing gain. However, it is impossible to mount multiple antennas on a small sensor node. To achieve MIMO gains in sensor networks, cooperative MIMO techniques have been proposed. These techniques allow multiple nodes to cooperate in signal transmission and/or reception. In [13], the authors analyze the diversity performance and propose corresponding space-time code designs for cooperative schemes involving a relay node. In [14], the energy efficiency and delay performance of cooperative MIMO techniques are analyzed for a single-hop system where it is shown that both energy and delay can be reduced within a certain transmission range. However, due to the energy and delay associated with the local information exchange within the cluster, the cooperative MIMO approach is less efficient than the traditional non-cooperative approach when the transmission distance between clusters is below some threshold.

In this paper, we combine the results in [11] and [14] to show how cooperative MIMO techniques can be applied to improve network performance. By jointly designing routing and link scheduling for networks composed of multiple clusters of nodes using cooperative MIMO, we show that the end-to-end performance can be dramatically improved. Moreover, our novel approach of distributed Alamouti coding provides diversity gain with no local information exchange, as is typically required in node cooperation [14].

The remainder of this paper is organized as follows. Section II describes the system model for the cooperative MIMO approach and proposes an equivalent SISO system to solve for the optimal routing and scheduling in the network. In Section III we analyze the delay performance and energy consumption of the proposed schemes with fixed link rates and compute the optimal energy-delay tradeoff curve. In Section IV we discuss the case where we allow link rate adaptation, which enables the full cross-layer optimization across routing, MAC, and link layers. Section V summarizes our conclusions.

II. System Model

We consider a sensor network composed of multiple clusters of nodes, as shown in Fig. 1. This figure shows clusters of nodes where the nodes within the same cluster are closely spaced and cooperate in signal transmission and/or reception. In this work we extend the cooperative strategy proposed in [14] to a multihop networking scenario, where we find the routing and scheduling that optimize energy and/or delay performance based on cooperative MIMO transmissions at each hop.

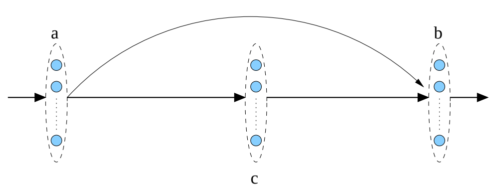

Fig. 1. A clustered network

We restrict our attention to the double-string network topology shown in Fig. 2, which represents regularly spaced sensors for data collection. In this topology there are clusters of two nodes, where within a cluster the nodes are separated by distance d_a while the distance between clusters is d_c with $d_c \gg d_a$. While Fig. 2 shows clusters of size $M = 2$, our design methodology

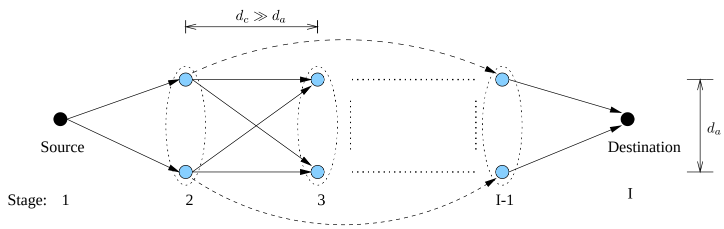

Fig. 2. A double-string network

applies to any cluster size. The highly regular topology of the double string network facilitates analysis, and also demonstrates potential performance gains for more general topologies. For the network in Fig. 2, there are $I-2$ stages of node clusters between the source and the destination. Thus, if the distance between the source and the destination is d, then the distance between the neighboring stages is $d_c = \frac{d}{I-1}$. We also assume that transmissions from stage m to stage n is allowed for any m and n with $1 \leq m \leq n \leq I$, where the source node is at stage 1 and the destination node is at stage I.

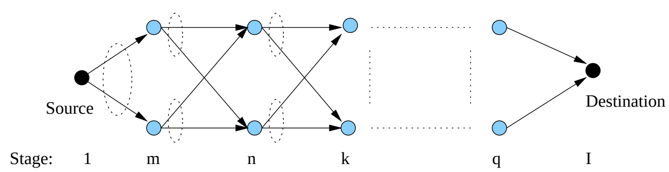

Fig. 3. Cooperative transmission

We assume that the source node generates data at L_1 packets per collection period T with a fixed packet size $v = 100$ bits. Therefore, the network needs to support a throughput of $S_0 = \frac{L_1}{T}$ packets per second (pps) between the source node and the destination node. We also assume a TDMA-based transmission scheme where the frame length is equal to T. Therefore, the network needs to convey L_1 packets from the source to the destination within each frame. We want to find a variable-length TDMA scheme where each transmission is assigned an optimal transmission time with the total sum bounded by T to minimize the energy consumed across the network within each frame. Due to the nature of TDMA, there is only one transmission in the network at any given time.

The nodes cooperate in the following manner. As shown in Fig. 3, within the first slot in each frame, the source node broadcasts a certain number of packets to the two nodes of the cluster at stage m, $2 \leq m \leq I$. If $m < I$, then the upper node at stage m acts as antenna 1 and the lower node acts as antenna 2. These antennas transmit two streams of codewords that are encoded according to a 2×1 Alamouti code [12]. Note that for a given time slot, the pair of nodes at stage m is allowed to transmit to any pair of nodes at stage n, $n > m$. The two nodes at stage n will decode the information independently and repeat the cooperative coding and transmission process. In addition, a pair of nodes may be assigned more than one time slot within each frame to transmit packets to different stages. Note that it is possible that the source node transmits all the packets directly to the destination node, if that is more efficient. The optimization of which clusters participate in the multihop routing and the corresponding transmission scheduling is performed off-line and communicated to all nodes prior to transmission. We also assume that the network is synchronized, which may be enabled by utilizing beacon signals in a separate control channel. Although the scheme just proposed is for node clusters of size $M = 2$, similar ideas can be applied to larger clusters.

For each link, we assume a flat Rayleigh fading channel, i.e., the channel gain between each transmitter and each receiver is a scalar. In addition, the mean path loss is modeled by a power falloff proportional to the distance squared, so the received power associated with transmission from stage m to stage n is given by

$$P_{mn}^r = \frac{P_{mn}^0}{G_0 d_{mn}^2} \tag{1}$$

where P_{mn}^0 is the transmit power at stage m, d_{mn} is the transmission distance between stage m and stage n, G_0 is the power attenuation factor at $d_{mn} = 1$ m, and P_{mn}^r is the received power [11]. We can also express the received power in terms of the received energy per bit as

$$P_{mn}^r = \bar{E}_b b_{mn} B, \tag{2}$$

where b_{mn} is the constellation size, B is the symbol rate which is approximately equal to the passband bandwidth, and $\bar{E}_b$ is the received energy per bit averaged over fading. Therefore, by combining Eq. (1) and Eq. (2), we can obtain the expression for the transmit power as

$$P_{mn}^0 = \bar{E}_b b_{mn} G_0 d_{mn}^2 B. \tag{3}$$

Note that we have $b_{mn} = \frac{W_{mn}\nu}{Bt_{mn}}$ where W_{mn} is the number of packets transmitted from stage m to stage n and t_{mn} is the transmission time. This relationship guarantees that all the $W_{mn}\nu$ bits can be sent from stage m to stage n within t_{mn} seconds. As long as B and ν are fixed, the value for any particular variable among b_{mn}, W_{mn}, and t_{mn} can be determined by the values of the other two variables.

As shown in [12], the instantaneous received SNR for a 2×1 Alamouti system is given by $\gamma_b = \frac{\|\mathbf{H}\|_F^2}{2}\frac{\bar{E}_b}{N_0}$ where $\mathbf{H} = [h_1\ h_2]$ with h_1 and h_2 zero mean circulant symmetric complex Gaussian (ZMCSCG) random variables with unit variance [12] and $N_0/2$ is the double-sided power spectral density for the AWGN noise. For the 2×1 MISO system with a constellation size b we can apply the Chernoff bound [16] to obtain the average probability of bit error as

$$\bar{P}_b \leq \frac{4}{b}\left(1 - \frac{1}{2^{\frac{b}{2}}}\right)\left(\frac{1.5\bar{E}_b b}{2N_0(2^b - 1)}\right)^{-2}, \quad b \geq 2, \tag{4}$$

from which we can derive an upper bound for $\bar{E}_b$ as shown below [14]:

$$\bar{E}_b \leq \frac{4}{3}\left(\frac{\bar{P}_b}{4}\right)^{-\frac{1}{2}} \frac{2^b - 1}{b^{\frac{1}{2}+1}} N_0.$$

By approximating this bound as an equality, we obtain an expression for $\bar{E}_b$, from which we can calculate P^0_{mn} according to Eq. (3).

The total power consumed in the transmitter power amplifier is given by [1]

$$P_t^{mn} = (1+\alpha)P^0_{mn}, \tag{5}$$

where α is defined by the power amplifier efficiency and other system parameters [1].

Therefore, the total power consumed in the two transmitter power amplifiers during the transmission from stage m to stage n is given by:

$$P_t^{mn} = \frac{4}{3}(1+\alpha)\left(\frac{\bar{P}_b}{4}\right)^{-\frac{1}{2}} \frac{2^{b_{mn}}-1}{\sqrt{b_{mn}}} N_0 G_0 d^2_{mn} B. \tag{6}$$

For QPSK with $b_{mn} = 2$, we have

$$P_t^{mn} = 2\sqrt{2}(1+\alpha)\left(\frac{\bar{P}_b}{4}\right)^{-\frac{1}{2}} N_0 G_0 d^2_{mn} B.$$

Therefore, the total power consumed during the transmission from stage m to stage n is given by

$$P_{mn} = P^m_{ct} + P^n_{cr} + P^{mn}_t, \tag{7}$$

where P^m_{ct} is the total transmitter circuit power consumption across stage m and P^n_{cr} is the total receiver circuit power consumption across stage n. Note that when $m = 1$, i.e., the SISO transmission is from the source node to other stages, P^{mn}_t is given by [15]

$$P_t^{mn} = (1+\alpha)\frac{1}{12\bar{P}_b}2^{b_{mn}} N_0 G_0 d^2_{mn} B. \tag{8}$$

However, after we know how to calculate P_{mn}, it is still difficult to incorporate the cooperative MIMO structure into the routing optimization model, which is addressed in [11] for the non-cooperative systems. Fortunately, we can apply a simple trick to make the problem manageable. Since all the transmissions occur between different pairs of nodes and the paring relationship is fixed, we can treat each pair of nodes in the same stage as one super node. Then the double-string network is simplified to a single-string network as shown in Fig 4, which can be treated as a virtual SISO system with the total number of nodes given by $N = I$. The total power required for transmission between two super nodes is given by Eq. (7). The corresponding energy or delay minimization problem can thus be modeled in the same way as in the SISO case, which will be discussed in the next section. For networks with an arbitrary cluster size M, similar equivalent SISO systems can be obtained with P_{mn} modified according to an $M \times 1$ MISO system.

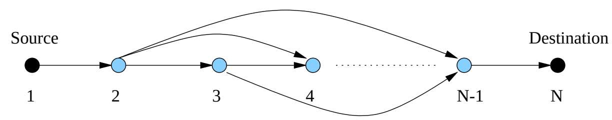

Fig. 4. Equivalent SISO system

A. Optimization of Equivalent SISO Systems with Arbitrary Link Rate

According to [11], for any network with one source node and one destination node, we can model the minimum-energy routing problem as an optimization problem when SISO transmissions are used for each link. The topologies shown in Fig. 2 (with $N = 2I - 2$) and Fig. 4 (with $N = I$) are special cases for such solvable networks if SISO transmissions are exclusively used. As in [11], we assume that the network is static such that the optimization can be done off-line before the network is deployed.

We now discuss the optimization model for SISO-based systems in detail. For node i, we use $\mathcal{N}_i$ to denote the set of nodes that send data to node i, and use $\mathcal{M}_i$ to denote the set of nodes that receive data from node i. We denote the normalized time slot length for the transmission over link $i \to j$ (from node i to node j) as $\delta_{ij} = \frac{t_{ij}}{T}$, where $\sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} \delta_{ij} \leq 1$. As introduced before, we use W_{ij} to denote the number of packets transmitted over link $i \to j$ during each period T. As discussed in [11] we assume three modes of operation for each node: active mode, sleep mode, and transient mode. To simplify the formulation we neglect the effect of the transient mode [1]. Thus, nodes i and j will be in active mode when link $i \to j$ is active, and will otherwise be in sleep mode where all the circuits are turned off to save energy. At node i, as introduced in [11], we use P_{ct}^i and P_{cr}^i to denote the circuit power consumption values for the transmitting circuits and the receiving circuits, respectively. The transmit power needed for satisfying a target probability of bit error P_b from node i to node j is denoted as P_t^{ij}. Therefore, the total average power spent on link $i \to j$ is given as

$$P_{ij} = \delta_{ij}(P_{cr}^j + P_{ct}^i + P_t^{ij})$$

and the total energy consumed over link $i \to j$ is given as $\epsilon_{ij} = TP_{ij}$.

As discussed in [11], to increase the network lifetime we can choose to minimize the total energy consumption as follows

$$\begin{array}{ll} \min & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} \epsilon_{ij} \\ \text{s. t.} & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} \frac{\nu W_{ij}}{B b_{ij}} \leq T \\ & \sum_{j \in \mathcal{M}_i} W_{ij} - \sum_{j \in \mathcal{N}_i} W_{ji} = L_i, \quad i = 1, \cdots, N \\ & 2 \leq b_{ij} \leq C_{ij}, \quad i = 1, \cdots, N-1, \quad j = 1, \cdots, N \end{array}, \tag{9}$$

where the first constraint is the TDMA constraint, the second constraint is the flow conservation constraint, which guarantees that at each node the difference between the total outgoing traffic and the total incoming traffic is equal to the traffic generated by the node itself, and C_{ij} in the third constraint is the maximum constellation size each link can use without violating certain peak power constraint. For the double-string information collection network topology shown in Fig. 4, we have $L_i = 0$, $i = 2, \cdots, N-1$, and $L_N = -L_1$ where the negative sign is due to the fact that the destination node has only incoming traffic.

Since W_{ij} and b_{ij} can only take integer values, the problem is an integer programming problem, which is not convex. Even if we allow these parameters to take on real values, the optimization problem is still not jointly convex over W_{ij} and b_{ij}. Fortunately, since we have the relationship $b_{ij} = \frac{\nu W_{ij}}{B t_{ij}}$, we can optimize over W_{ij} and t_{ij} instead. As such, the problem becomes jointly convex over t_{ij} and W_{ij}. The convexity proof can be boiled down to proving the convexity of two special functions, which is shown in Appendix A and B. Finally, the optimization problem becomes

$$\begin{array}{ll} \min & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} \epsilon_{ij} \\ \text{s. t.} & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} t_{ij} \leq T \\ & \sum_{j \in \mathcal{M}_i} W_{ij} - \sum_{j \in \mathcal{N}_i} W_{ji} = L_i, \quad i = 1, \cdots, N \\ & \frac{\nu W_{ij}}{C_{ij} B} \leq t_{ij} \leq \frac{\nu W_{ij}}{2B}, \quad j \in \mathcal{M}_i, \quad i = 1, \cdots, N-1 \end{array}, \tag{10}$$

which is convex over t_{ij} and W_{ij} if we allow them to take real values. To reduce the relative error caused by the relaxation, we can use integer programming techniques such as the Branch and Bound algorithm [17], which is discussed in [11] in more detail.

Different scheduling (ordering) of the optimal time slot assignments, the t_{ij}'s, will lead to different delay performance, although they all have the same energy efficiency. It is shown in [11] that the minimum packet delay among all possible schedules is equal to the frame length T, and a simple algorithm exists to find such a minimum-delay schedule for any loop-free network with one sink node. Thus, by solving the problem in Eq. (10), we can find the minimum possible energy required to transfer a given number of packets within a delay deadline T. Alternatively, instead of minimizing energy under a delay constraint, we can also consider the dual problem of minimizing delay under an energy constraint. Specifically, given a total energy budget E_M per period, we can find the minimum possible value for $T = \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} t_{ij}$ that is required to finish the transfer of a given number of packets. The dual problem is characterized as

$$\begin{array}{ll} \min. & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} t_{ij} \\ \text{s. t.} & \sum_{j \in \mathcal{M}_i} W_{ij} - \sum_{j \in \mathcal{N}_i} W_{ji} = L_i, \quad i = 1, \cdots, N \\ & \sum_{i=1}^{N-1} \sum_{j \in \mathcal{M}_i} \epsilon_{ij} \leq E_M \\ & t_{ij} \geq 0, \quad j \in \mathcal{M}_i, \quad i = 1, \cdots, N-1 \end{array}. \tag{11}$$

The optimal t_{ij}'s given by solving Eq. (10) and Eq. (11) can take arbitrary real values. Thus, the resulting variable length TDMA scheme is impractical, since it will require an infinite number of bits to describe the time slot assignment. To alleviate this problem, we can divide the frame into unit slots with length Δ. After we obtain the optimal values for the $t'_{ij}s$, the optimal number of unit slots assigned to each link is given by rounding $\frac{t_{ij}}{\Delta}$ to the nearest integer. As long as Δ is small enough, the performance degradation caused by the rounding is negligible. Thus, in this paper we just focus on finding the optimal real-valued t_{ij}'s.

B. Optimization of Equivalent SISO Systems with Fixed Link Rate (QPSK)

For the simple case where we fix the transmission rate, the optimization problem discussed in the last section can be simplified to a Linear Programming (LP) problem, which can be efficiently solved [18]. Specifically, if we use QPSK transmissions for all the links, the optimization problems shown in Eq. (10) and Eq. (11) can be rewritten as

$$\begin{array}{ll} \min & T\sum_{i=1}^{N-1}\sum_{\mathcal{M}_i} P_{ij} \\ \text{s. t.} & \sum_{i=1}^{N-1}\sum_{j\in\mathcal{M}_i} t_{ij} \leq T \\ & \sum_{j\in\mathcal{M}_i} t_{ij} - \sum_{j\in\mathcal{N}_i} t_{ji} = \frac{L_i}{S_a}, \quad i=1,\cdots,N \\ & t_{ij} \geq 0, \quad j\in\mathcal{M}_i, \quad i=1,\cdots,N-1 \end{array}, \qquad (12)$$

where $S_a = \frac{B}{50}$ pps for QPSK and

$$\begin{array}{ll} \min. & \sum_{i=1}^{N-1}\sum_{j\in\mathcal{M}_i} t_{ij} \\ \text{s. t.} & \sum_{j\in\mathcal{M}_i} t_{ij} - \sum_{j\in\mathcal{N}_i} t_{ji} = \frac{L_i}{S_a}, \quad i=1,\cdots,N \\ & \sum_{i=1}^{N-1}\sum_{j\in\mathcal{M}_i} P_{ij} t_{ij} \leq E_M \\ & t_{ij} \geq 0, \quad j\in\mathcal{M}_i, \quad i=1,\cdots,N-1 \end{array}, \qquad (13)$$

respectively, where we see both the objective function and the constraints are linear over the design variable t_{ij}'s.

III. ENERGY-DELAY TRADEOFF WITH FIXED LINK RATES

In the last section we introduced the optimization models to minimize energy subject to a constraint in delay or to minimize delay subject to a constraint in energy. In this section, we provide numerical results for these optimizations, along with the optimal energy-delay tradeoff curves. These results illustrate the performance benefit of cross-layer design.

We start with the case where we assume that all the nodes support a fixed transmission rate. Specifically, we assume QPSK transmissions with a $B = 10$ KHz symbol rate. The packet transmission rate (denoted as S_a) at each node is given by $S_a = 200$ pps. By fixing the link rate, we simplify the cross-layer design model to consider only the routing and MAC layers. Since the constellation size b is fixed, the design variables are t_{ij}'s, over which the optimization problems described in Eq. (12) and Eq. (13) are all Linear Programming (LP) problems.

For a network where each link has a fixed transmission rate, multihop transmission consumes less total transmission power than single-hop transmission as long as the path loss is proportional to $\frac{1}{d^\kappa}$ with $\kappa > 1$. This is true for both non-cooperative and cooperative MIMO systems. However, when the delay constraint is tight, multihop transmissions may not be feasible since the total delay is monotonically increasing with the number of hops [11]. In addition, when circuit energy consumption is considered, as shown in [11], multihop transmissions may not be more energy efficient than single-hop transmissions since the relay nodes consume extra circuit processing energy. By solving the optimization problems given in Eq. (12) we can determine when multihop transmissions should be utilized to minimize energy consumption.

For our numerical results we consider a double-string network with ten stages ($I = 10$), $d = 270$ m, $S_a = 200$ pps, and $L_1 = 60$ packets. As in [14], we take $\bar{P}_b = 10^{-3}$, $G_0 = 30$ dB, $\alpha \approx 7.6$, and $N_0 = -134$ dBm/Hz. For both the non-cooperative and cooperative MIMO systems, if the frame length $T \leq \frac{L_1}{S_a} = 0.3$ s, single-hop transmission is the only option since the frame length T is not large enough for multiple hops to take place. When $T > 0.3$ s, we have the option to use multihop routing to save transmission energy. The minimum energy transmission schemes with $T = 1.5$ s for the non-cooperative and cooperative MIMO systems are shown in Fig. 5 and Fig. 6, respectively. These figures show the optimal routing when only transmission energy is considered or when both circuit and transmission energy is included. The number beside each link is the optimal time slot length assigned to that link. For both systems, we see that when circuit energy consumption is included, energy efficient transmissions involve a fewer number of hops than when only transmission energy is considered. Note that in Fig. 6, only the eight intermediate super nodes are shown, which represent clusters of two nodes. The transmissions in this figure represent cooperative MIMO transmissions. In addition, as proved in [11], we can always find an optimal transmission order for all the active links to guarantee that all the L_1 packets arrive at the destination node within the current frame. Therefore, we call the frame length T the scheduling delay.

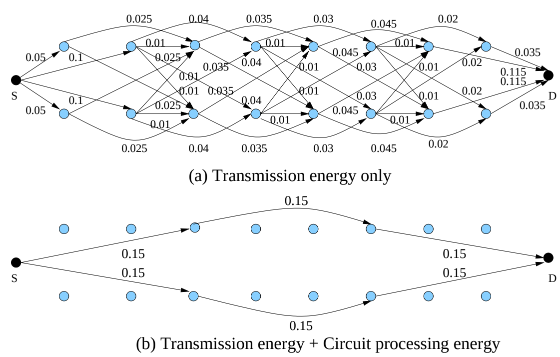

Fig. 5. Minimum-energy routing and scheduling (SISO-based)

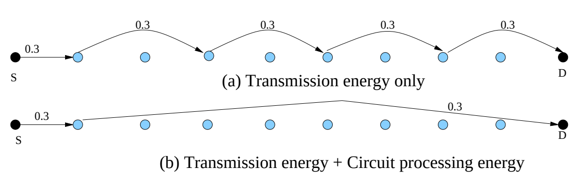

Fig. 6. Minimum-energy routing and scheduling (MISO-based)

For a given network topology, the achievable energy-delay region consists of all the achievable energy-delay pairs. The energy-delay region is a convex set. This is because if energy-delay points (ϵ_1, T_1) and (ϵ_2, T_2) are contained in the energy-delay region, then any convex combination of these points can be achieved by time-sharing between the transmission schemes corresponding to the two end points. Hence, any convex combination of these points are contained in the achievable energy-delay region. Here, we calculate the Pareto-optimal energy-delay tradeoff which characterizes the minimum possible delay for a given energy consumption (or vice versa), and the optimal tradeoff curve defines the boundary of the achievable energy-delay region.

The optimal tradeoff curve can be found by varying the value of β in the following optimization problem.

$$\begin{array}{ll} \text{min.} & \sum_{i=1}^{N-1}\sum_{j\in\mathcal{M}_i} t_{ij} + \beta\sum_{i=1}^{N-1}\sum_{j\in\mathcal{M}_i} P_{ij}t_{ij} \\ \text{s. t.} & \sum_{j\in\mathcal{M}_i} t_{ij} - \sum_{j\in\mathcal{N}_i} t_{ji} = \frac{L_i}{S_a}, \quad i = 1,\cdots,N \\ & t_{ij} \geq 0, \quad j\in\mathcal{M}_i, \quad i=1,\cdots,N-1 \end{array}, \tag{14}$$

where the first term in the objective function is the delay and the second term is the total energy consumption weighted by a scanning parameter β. The resulting problem is a LP problem for each β when the link rate is fixed, which can be efficiently solved using existing techniques [18].

To give a numerical example, we consider the same ten-stage double-string network with the same system parameters as we used for obtaining the results in Fig. 5 and Fig. 6. The optimal energy-delay tradeoff curves for both the non-cooperative and cooperative MIMO systems are shown in Fig. 7 and Fig. 8 , where we see that the optimal curve for the cooperative MIMO system (labeled as coop MIMO) is strictly below that of the non-cooperative system (labeled as non-coop), which means cooperative MIMO schemes can reduce both energy and delay. For both the case where the circuit processing energy is included and the case where it is not included, we see that the two curves converge to the same point on the far right of the curves, which corresponds to the scenario where the source node transmits all the packets directly to the destination node. This is expected since delay is minimized by single-hop transmissions.

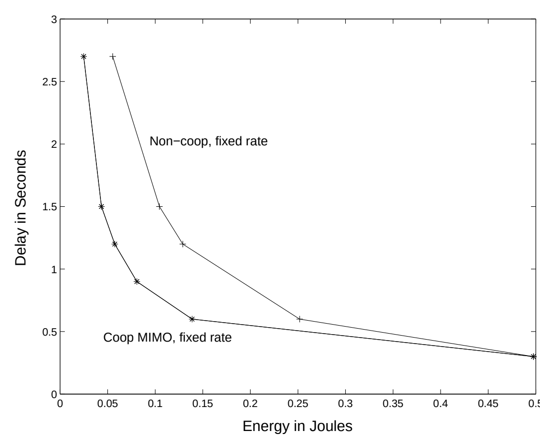

Fig. 7. Transmission Energy only

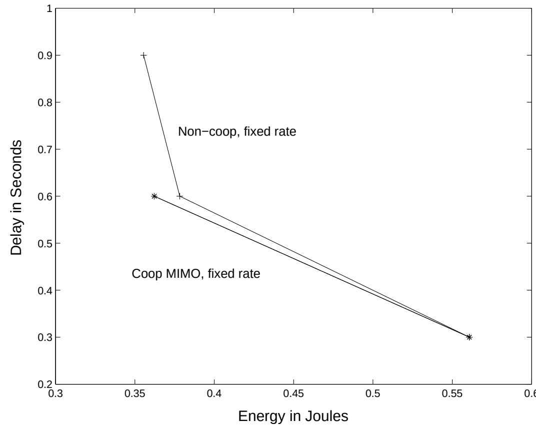

Fig. 8. Circuit energy included

IV. Delay-Energy Tradeoff with Link Rate Adaptation

In the last section, we have shown that cooperative MIMO can reduce both energy and delay, even though the link rate is fixed. We now investigate what further performance gains can be obtained by allowing the link rate to be optimally chosen according to the transmission distances. By adding link rate adaptation, we extend our cross-layer optimization over the link, MAC, and network layers.

With the flexibility of rate adaptation, the routing delay can be reduced, since we can always use higher constellation sizes to reduce the transmission time for links that carry more traffic. Moreover, we can reduce circuit energy consumption in the relay nodes by assigning higher constellation sizes to the links on the particular multihop route, since a higher constellation size means a shorter transmission time, which is also the circuit active time in the relay nodes. These benefits give the routing layer more freedom to choose the optimal route. We now illustrate the performance gain achieved with link layer adaptation by deriving the optimal energy-delay tradeoff curve.

The optimal energy-delay tradeoff curves with link rate adaptation can be obtained by using the same model as in Eq. (14). The difference from the fixed-rate case is that the value for the function P_{ij} is now defined by Eq. (6-8). As a result, the design variables for the optimization problem become W_{ij}'s and t_{ij}'s. The problem can still be solved using convex optimization techniques as we discussed for the model in Eq. (10).

For our numerical results, we use the same network example as in Section III. For the case where we only consider the transmission energy, the optimal energy-delay tradeoff curve is shown in Fig. 9, where we see that the benefit of rate adaptation is not obvious except that the delay can be further reduced at the expense of energy. On the left side of point A, the two curves for the cooperative MIMO systems have almost merged due to the fact that QPSK is used in both systems to minimize the transmission energy. The slight difference between the two curves on the left side of point A is just due to some numerical rounding errors.

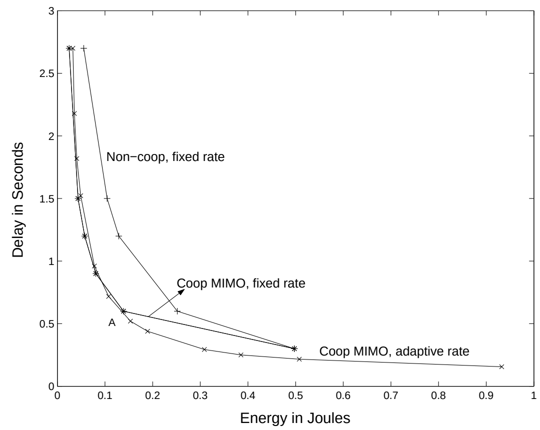

Fig. 9. Transmission Energy only

For the case where we consider both the transmission energy and the circuit processing energy, the optimal energy-delay tradeoff curve is shown in Fig. 10, where we see dramatic performance improvement achieved by the cooperative MIMO system with rate adaptation, since rate adaptation can always minimize the sum of the transmission energy and the circuit processing energy and gives the upper layers more freedom to choose optimal multihop routes.

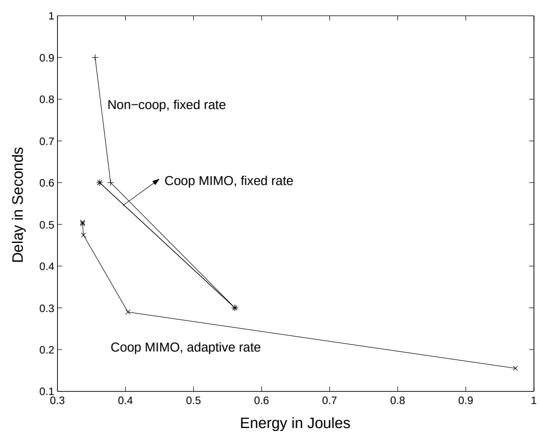

Fig. 10. Circuit energy included

For the double-string topology we consider here, the traffic is evenly distributed across the network. For networks with unevenly distributed traffic, we can qualitatively describe another potential benefit of adding the link layer adaptation to the cross-layer design. Consider an arbitrary network with a single destination node, as shown in Fig. 11. For such networks, a large amount of traffic from the source nodes is typically routed through the nodes surrounding the destination node. Thus, the neighborhood of the destination is a heavy traffic region. The links in this heavy traffic region need to support high transmission rates (corresponding to large constellation sizes for MQAM with a fixed symbol rate), which requires a high transmit power. Assume a TDMA MAC protocol. If each link can adaptively choose its rate, then the links in the light traffic region can transmit with a higher constellation size to reduce the required transmission time. The saved time slots can be reassigned to the links in the heavy traffic region. Since the links under heavy traffic now have a longer available transmission time, they can transmit information with lower constellation sizes to save transmission energy without jeopardizing the average throughput. By this simple interaction, the overall network energy consumption can be reduced.

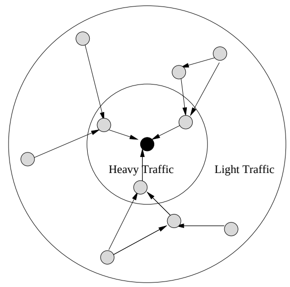

Fig. 11. A network example

V. Conclusions

We show that cooperative MIMO coupled with cross-layer optimization can significantly improve the energy-delay tradeoff in wireless networks. If the cooperation is properly executed and jointly designed with upper layers, no local information exchange between cooperating nodes is needed, and the optimal routing and transmission schemes can be found using convex optimization techniques. We provide numerical examples demonstrating the performance improvements of cooperative MIMO over non-cooperative methods. The performance difference is especially dramatic when rate adaptation is allowed in the link layer.

Appendix

A. Proof of convexity for ϵ_{ij} over $t_{ij} > 0$, $W_{ij} > 0$, and $\frac{W_{ij}\nu}{Bt_{ij}} \geq 2$ (for Cooperative MIMO links)

Proving the convexity of the function (over W_{ij} and t_{ij})

$$\epsilon_{ij} = x_{ij} \frac{\left(2^{\frac{\nu W_{ij}}{Bt_{ij}}} - 1\right)}{\left(\frac{\nu W_{ij}}{Bt_{ij}}\right)^{3/2}} \nu W_{ij} + y_{ij} t_{ij}$$

where x_{ij} and y_{ij} are some system constants, is equivalent to proving the convexity of the function

$$f(W,t) = \frac{2^{\frac{W}{t}} - 1}{\sqrt{W}} t^{3/2}$$

over W and t, after we remove all the linear terms and redefine $W = \frac{W_{ij}\nu}{B}$ and $t = t_{ij}$. For function $f(W,t)$, the Hessian matrix is given by

$$\mathbf{H} = \left[\begin{array}{cc} a(W,t) & b(W,t) \\ b(W,t) & c(W,t) \end{array} \right],$$

where we have

$$\begin{aligned} a(W,t) &= \frac{2^{W/t}W4\ln 2(W\ln 2 - t) + 3t^2(2^{W/t}-1)}{4t^{1/2}W^{5/2}}; \\ b(W,t) &= -\frac{2^{W/t}W4\ln 2(W\ln 2 - t) + 3t^2(2^{W/t}-1)}{4t^{3/2}W^{3/2}}; \\ c(W,t) &= \frac{2^{W/t}W4\ln 2(W\ln 2 - t) + 3t^2(2^{W/t}-1)}{4t^{5/2}W^{1/2})}. \end{aligned}$$

We can show that $a(W,t) > 0$ when $\frac{W}{t} = b \geq 2$ due to the fact that its denominator as well as the first term and the second term in the numerator are all strictly positive. We can further show that $b^2(W,t)/a(W,t) - c(W,t) = 0$. For a matrix in the form of $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ with $a > 0$, it is positive semi-definite as long as we have $b^2a^{-1} - c = 0$ according to Schur's complement condition [18]. Therefore, we can claim that $\mathbf{H}$ is positive semi-definite or equivalently, $f(W,t)$ is convex over W and t.

B. Proof of convexity for ϵ_{ij} over $t_{ij} > 0$, $W_{ij} > 0$, and $\frac{W_{ij}\nu}{Bt_{ij}} \geq 2$ (for SISO links)

Proving the convexity of the function (over W_{ij} and t_{ij})

$$\epsilon_{ij} = x_{ij}\frac{2^{\frac{\nu W_{ij}}{Bt_{ij}}}}{\frac{\nu W_{ij}}{Bt_{ij}}}\nu W_{ij} + y_{ij}t_{ij}$$

where x_{ij} and y_{ij} are some system constants, is equivalent to proving the convexity of the function

$$f(W,t) = 2^{\frac{W}{t}}t$$

over W and t, after we remove all the linear terms and redefine $W = \frac{W_{ij}\nu}{B}$ and $t = t_{ij}$. For function $f(W,t)$, the Hessian matrix is given by

$$\mathbf{H} = \begin{bmatrix} a(W,t) & b(W,t) \\ b(W,t) & c(W,t) \end{bmatrix},$$

where we have

$$\begin{aligned} a(W,t) &= \frac{2^{W/t}}{t}\ln^2 2; \\ b(W,t) &= \frac{-W2^{W/t}}{t^2}\ln^2 2; \\ c(W,t) &= \frac{W^2 2^{W/t}}{t^3}\ln^2 2. \end{aligned}$$

Following the same argument as in Appendix A, we can show $f(W,t)$ is convex.

REFERENCES

[1] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," to appear in *IEEE Trans. Wireless Commun.*, September, 2005. Also available at http://wsl.stanford.edu/Publications.html

[2] A. J. Goldsmith and S. B. Wicker, "Design challenges for energy-constrained ad hoc wireless networks," *IEEE Wireless Communications Magazine*, pp. 8-27 Aug. 2002.

[3] T. ElBatt and A. Ephremides, "Joint Scheduling and Power Control for Wireless Ad Hoc Networks,", *IEEE Transactions on Wireless Communications*, Vol. 1, pp. 74-85, 2004.

[4] R. L. Cruz and A. V. Santhanam, "Optimal Routing, Link Scheduling and Power Control in Multi-hop Wireless Networks," Proceedings of *IEEE INFOCOM*, 2003.

[5] U. C. Kozat, I. Koutsopoulos, and L. Tassiulas, "A Framework for Cross-layer Design of Energy-efficient communication with QoS Provisioning in Multi-hop Wireless Networks," proceedings of *IEEE INFOCOM*, 2004.
[6] R. Bhatia and M. Kodialam, "On Power Efficient Communication over Multi-hop Wireless Networks: Joint Routing, Scheduling and Power Control," Proceedings of *IEEE INFOCOM*, 2004.
[7] L. Xiao, M. Johansson, and S. Boyd, " Simultaneous Routing and Power Allocation in CDMA Wireless Data Networks," to appear at *IEEE Trans. on Communications*, 2004. Also available at http://www.stanford.edu/∼boyd/srra.html
[8] C. Schurgers, O. Aberthorne, and M. B. Srivastava, "Modulation scaling for energy aware communication systems," *International Symposium on Low Power Electronics and Design*, pp. 96-99, 2001.
[9] R. Min and A. Chadrakasan, "A framework for energy-scalable communication in high-density wireless networks," *International Syposium on Low Power Electronics Design*, 2002.
[10] S. Cui, A. J. Goldsmith, and A. Bahai, "Joint modulation and multiple access optimization under energy constarints," Proceedings of *Globecom'04*, 2004.
[11] S. Cui, R. Madan, S. Lall, and A. J. Goldsmith, "Energy Minimization and Delay Analysis in TDMA-based Sensor Networks," submitted to *IEEE Trans. Wireless Commun.*, 2004. Also available at http://wsl.stanford.edu/Publications.html
[12] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, preprint, Cambridge University Press, Cambridge, UK, 2003.
[13] R. U. Nabar, H. Bolcskei, and F. W. Kneubuhler "Fading relay channels: performance limits and space-time signal design," *IEEE J. Selet. Areas Commun.*, pp. 1099-1109, Vol. 22, No. 6, August, 2004.
[14] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and Cooperative MIMO in Sensor Networks," *IEEE J. Select. Areas of Commun.*, pp. 1089-1098, Vol. 22, No. 6, August, 2004.
[15] S. Cui, A. J. Goldsmith, and A. Bahai, "Modulation Optimization under Energy Constraints," *Proceedings of ICC'03*, Alaska, May, 2003.
[16] J. G. Proakis, *Digital Communications*, 4th Ed. New York: McGraw-Hill, 2000.
[17] R. S. Garfinkel and G. L. Nemhauser, *Interger Programming*, John Wiley & Sons, New York, 1972.
[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge Univ. Press, Cambridge, U.K., 2003.