Integrated Design of STBC-based Virtual-MIMO and Distributed Compression in Energy-Limited Wireless Sensor Networks

Madhavi L. Chebolu and Sudharman K. Jayaweera
Department of ECE, Wichita State University, Wichita, KS 67260, USA.
Email: sudharman.jayaweera@wichita.edu

***Abstract*— An integrated design of a joint distributed source coding and an adaptive signal processing scheme for a virtual multiple input multiple output (V-MIMO) communications-based, cooperative wireless sensor network is proposed. We employ a distributed coding scheme that exploits inherent correlations among sensor data and an adaptive correlation tracking technique based on Recursive Least Squares (RLS) algorithm that provides improved energy efficiency compared to previously proposed such joint coding and signal processing schemes for single input single output (SISO) communications based systems. An integration scheme with virtual space-time block coding based M-ary quadrature amplitude modulation (M-QAM) communication is also developed. The system efficiency is analyzed for various transmission distances and various channel conditions. Our results show that the proposed system achieves significant energy savings compared to conventional designs beyond certain transmission distances. The achieved energy efficiencies and low decoding error rates justify the application of the proposed scheme in energy constrained wireless sensor networks.**

I. INTRODUCTION

There is a new-found interest in research geared towards developing large-scale, low-cost wireless sensor networks. Some of the driving applications of such sensor networks include environmental observation and forecasting, habitat monitoring, remote ecological monitoring, biomedical applications such as health monitoring as well as military applications. Most of these networks consist of battery-operated sensors which introduce strict energy constraints. Thus, energy efficiency is of prime importance in the design of such wireless sensor networks. Previously proposed techniques for energy conservation in wireless sensor networks include, for example, information packet aggregation along sensor communication paths to reduce header overheads and switching off nodes when they are not functioning [1].

Multiple Input Multiple Output (MIMO) systems have found great importance in wireless communications due to possible enormous performance improvements over SISO systems [2]–[4] and [5]. Various MIMO techniques like Space Time Block Codes (STBC), Vertical Bell Laboratories Layered Space Time (V-BLAST), layered space-time designs and smart antenna techniques [6]–[10], have already been proposed for wireless Local Area Network (LAN) and cellular systems. According to [11], MIMO systems have been shown to outperform systems with receiver diversity alone in the case of interference-limited cellular systems. But the main drawbacks of MIMO techniques are that they could require complex receiver circuitry, signal processing as well as large physical dimensions to accommodate multiple antennas. But, in most wireless sensor networks, nodes are battery-operated and also cannot accommodate multiple antennas because of their physical size limitations making direct application of MIMO techniques unsuitable for sensor networks. However the recently proposed Virtual MIMO (V-MIMO) concept allows the realization of MIMO communication in wireless sensor networks via so-called local communications among sensors [12].

A scheme similar to virtual MIMO has also been investigated under the term VAA (Virtual Antenna Arrays) in [13]. There, antenna diversity at mobile terminals (each with one antenna element) is achieved through VAA-space time block codes. Specifically 2×2 VAA is illustrated, where a base station (BS) transmits the symbols in a sequence similar to that of Alamouti [6] to distributed mobile terminals. One of the mobile terminals acting as a relaying mobile station (MS), relays its received data to the target MS over another fading channel (with SISO implementation). At target MS, the combining of all received symbols (a modified version of Alamouti combining) is performed to retrieve the originally transmitted symbols of BS. In relation to the above, higher order space-time codes for VAA are also addressed in [14]. VAA application to wide-band code division multiple access (W-CDMA) and the effect of different configurations in a wireless mobile system with VAA are demonstrated in [15] and [16] respectively. Information theoretic results characterizing the capacity for proposed VAA have been presented in [17]. Resource allocation strategies have also been derived where the analysis is performed exclusively for space-time block encoded transceivers.

Till now, the research on virtual MIMO for wireless sensor networks have primarily been limited to demonstration of energy efficiencies compared to traditional SISO systems. The integration of virtual MIMO with other energy saving techniques such as distributed compression and/or signal processing has not been addressed at all. While virtual MIMO itself could provide significant communication energy savings, it is also important to design efficient integration techniques that lead to improved system performance. In this paper we take an initial step towards such an integrated design of virtual MIMO based wireless sensor systems. In particular, we integrate distributed coding and adaptive signal processing with V-MIMO architecture.

Our approach to energy efficiency in wireless sensor networks is based on the assumption that the data from different sensors exhibit spatio-temporal correlations. A scheme exploiting the redundancy caused by these inherent correlations to compress sensor data before transmission was proposed in [18]. This scheme can be applied to network models having two types of nodes: a data gathering node (DGN) and sensing or data collection nodes, where the DGN is assumed to be less energy-constrained relative to the sensing nodes. Thus, data gathering node performs most of the complex computations, based on which the sensors compress their data before transmitting them to the DGN. In [18] and [19], a Least Mean Squares (LMS) based correlation tracking algorithm was proposed. However, an RLS-based adaptive algorithm can provide a much faster rate of convergence and accurate estimation compared to LMS leading to better compression rates resulting in fewer decoding errors and increased energy savings [20]. Since DGN is assumed to have no (or at least less) energy constraint, it is reasonable to assume that it can support the somewhat increased complexity of RLS algorithm.

We demonstrate that with a judicious integration of such distributed compression techniques with the V-MIMO architecture, both the number of required transmissions from data collection sensors as well as energy per each transmission can be reduced. Throughout this paper, we assume an M-ary Quadrature Amplitude Modulation (M-QAM) for communication. The proposed system's performance is compared to various reference systems. These systems include V-MIMO without distributed compression, SISO without compression, and SISO with compression. These comparisons have been made with a different channel conditions over different transmission distances. Our results establish that energy efficiencies are significantly improved in all cases beyond certain transmission distances. In addition to that, the proposed V-MIMO based system can also lead to a significant reduction in decoding errors compared to that of reference systems.

The remainder of this paper is organized as follows: In Section II the virtual MIMO architecture is introduced along with the related energy expenditure expressions. Next, the proposed integrated scheme is detailed followed by the derivation of energy consumption expressions for the complete system. The test-bed used to analyze the energy-efficiency of the proposed scheme and the experimental results are given in Section III. Finally Section IV summarizes the conclusions of this work and presents possible future work.

II. INTEGRATED DESIGN OF V-STBC AND DISTRIBUTED COMPRESSION FOR WIRELESS SENSOR NETWORKS

A. System model

Consider a wireless sensor network consisting of low-end data collection sensor nodes and a high-end DGN. The DGN collects and processes the data from all low-end nodes, and also provides feedback to the sensors if needed. Although several V-MIMO schemes have been proposed for sensor networks [12], [21], in this paper we concentrate on space-time block coding based communication [6]. In order to realize MIMO communications in a sensor network, the sensor nodes should be able to accommodate multiple antennas. But, the sensor nodes can be of small dimensions making it difficult to employ multiple antennas. Thus, in [21], [22] sensor cooperation is introduced to achieve the so-called *virtual MIMO communications*.

Consider a set of sensors in a network that has data to be transmitted to the DGN. A group of nodes that are close to each other forms a cluster as in Fig. 1. Each node associates itself with at least one pre-determined cluster. Information sharing among sensor nodes within a cluster (called local communication) is required for a cluster

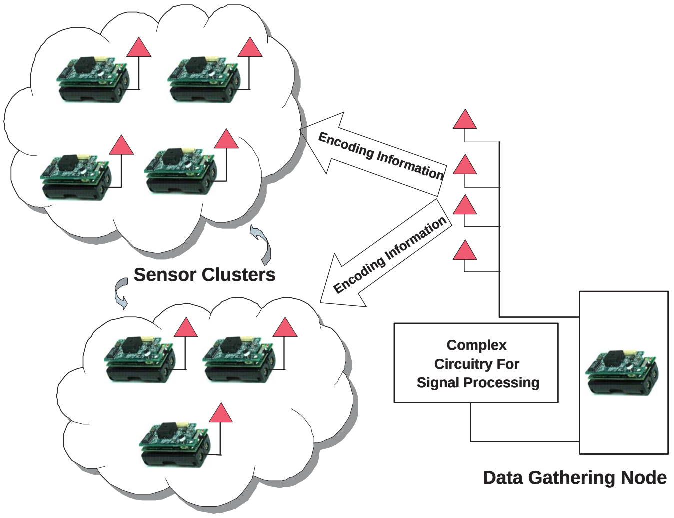

Fig. 1. Sensor Cluster Formation and Encoding Information Transmission by DGN

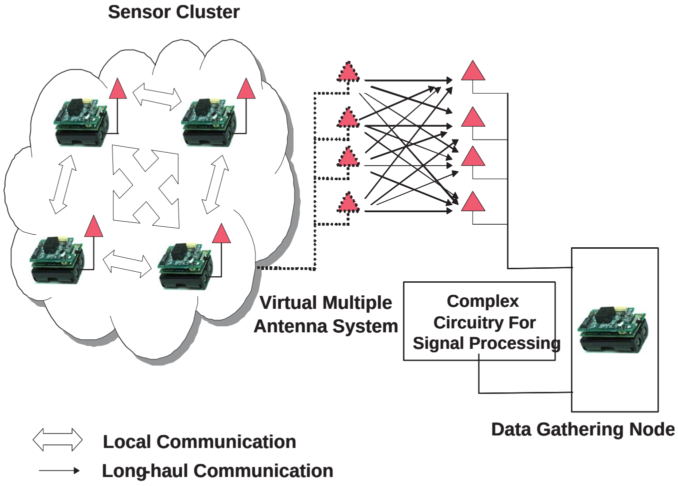

Fig. 2. Local and Long-haul Communications

to form a virtual multiple antenna system where each sensor node acts as an antenna element of a centralized antenna array, which can be seen in Fig. 2. The local communication within the cluster can be implemented, for example, using time division multiplexing. Receiver side (DGN side) inter-sensor communications can also be implemented if the DGN may be able to accommodate only one or two antennas to utilize the MIMO capability. The proposed communication strategy of [13] can be performed in sensor networks where one sensor cluster to another sensor cluster communication or sensor cluster - DGN communication is necessary, of course at the cost of relaying transmission.

This virtual multiple antenna system enables STBC coding of sensor data from a particular cluster to the DGN. Note that unlike sensor nodes, the DGN may be able to accommodate multiple antennas. This assumption allows us to achieve true MIMO capability with local communications only on the transmitter side inside a cluster and creating a set of virtual dual antenna array systems in the network. Transmission of STBC sensor data from a virtual antenna array to the DGN is known as long-haul communication (Figure 2 shows both local and long-haul communication process).

For convenience, consider a network with each cluster having two sensor nodes and a DGN. Let one of the sensors (reference sensor) always send its data uncoded or compressed with respect to its past readings. The other sensor can then compress its data with respect to its past readings and that of the reference sensor's readings, based on the sensor correlation information provided by DGN as in Fig. 1. Thus, each sensor can compress its data without any inter-sensor communication. But for a cluster to act as a multiple antenna unit, inter-sensor communication inside the cluster is necessary after compression of the original data. Though, after local communication each sensor of a cluster has data from other sensors, it cannot (and does not need to) determine the correlation structure required for data compression since the data shared is already compressed using the side information from the DGN.

The DGN collects the information from a cluster and computes the data correlations, based on which sensors compress their data. We assume that each sensor of a cluster receives the encoding information from DGN periodically every T_{Rec} time period. Each sensor compresses all samples within the entire period of T_{Rec} (which we denote by N_{Rec}) and then combines them into a single stream before the local communication. Thus, during correlation tracking the local communications are performed once in every T_{Rec} time period. As the encoding information of each sensor is broadcasted by the DGN, every node has knowledge of the encoding information of all the other sensors in the cluster. As the local communications are performed with compressed data, the transmission requirements are reduced.

Different digital modulation schemes can be used for both local and long-haul communications. Note that, with the assumed M-QAM modulation it may be required to add few extra bits (called don't-care bits) to the information bit stream to make the total number of bits for modulation divisible by b ($= \log_2 M$). If M-QAM is used for local communication, each sensor node can retrieve the originally transmitted information after demodulation, simply by removing the extra don't-care bits. The number of don't-care bits (N_{DC}) added will depend on the corresponding encoding information of a particular sensor. Note that, if transmission is performed at each and every sample instant, the don't-care bits need to be added every time. But, if a block transmission is done once every T_{Rec}, as proposed here, the number of don't-care bits to be added is relatively small compared to the total number of bits in a stream thereby reducing the extra energy consumption.

After local communications, each sensor demodulates the received data symbols broadcasted by other sensors of the cluster and combines them to form a single stream of data. Therefore sensors in a cluster act as antenna elements of a multiple transmitting antenna system. Assuming that the data streams available to all the antenna elements are the same, STBC is implemented for long-haul communication. Let us denote by $b_k^{(j)}$ the symbol stream of the j-th sensor and for simplicity assume that $N_s = 2$ (N_s is the number of sensors within a cluster). After the local communication inside the cluster, the two nodes have following symbols at time k : (i) $b_k^{(1)}$ and $\hat{b}_k^{(2,1)}$ at node 1 (ii) $b_k^{(2)}$ and $\hat{b}_k^{(1,2)}$ at node 2, where $\hat{b}_k^{(j,j')}$ is the estimate of $b_k^{(j)}$ at node j'. Then, the space-time block coding scheme (the Alamouti scheme) is implemented as follows: At the first time instant nodes 1 and 2 transmit $b_k^{(1)}$ and $b_k^{(2)}$, respectively. At the second time instant nodes 1 and 2 transmit $-\left(\hat{b}_k^{(2,1)}\right)^*$ and $\left(\hat{b}_k^{(1,2)}\right)^*$ respectively (see Fig. 3). Of course error propagation will occur if any errors occurred during the local communications. However, one can ensure that the error rate during the local communications is below a certain level by judicial choice of system parameters (note that the clusters have a very small radius so that the transmission power required to ensure a good error rate can be relatively small). Figure 3 illustrates the steps involved in the implementation of V-STBC communications (local and long-haul) for a 2 × 2 V-MIMO case along with distributed compression after the reception of encoding information from DGN.

In [21] M-QAM constellation optimization with respect to transmission distance was also performed to improve the system efficiency. We may also consider such optimum rates in our system in order to further improve the energy efficiency. However, distant-dependent, variable-rate communication may not be realistic in a low-power wireless sensor network. Thus, in our numerical results we consider fixed-rate M-QAM systems to demonstrate the realistic performance gains of the proposed system.

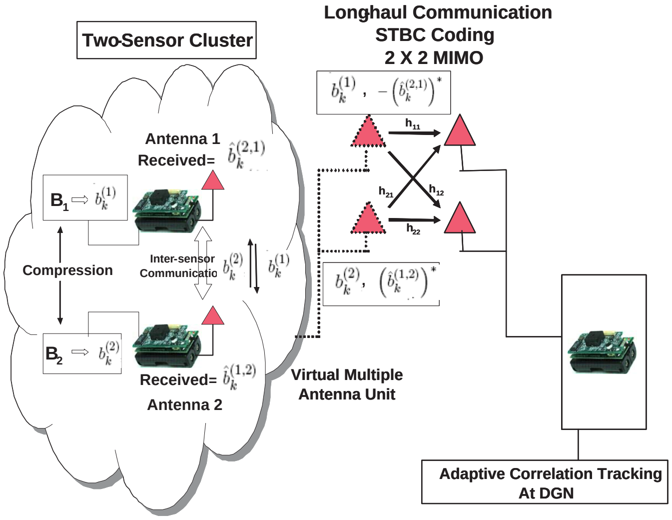

Fig. 3. Example of a Two-Sensor Cluster and DGN

In this paper, we consider an AWGN channel for local communications and two-sensor clusters and a DGN with either one or two antenna elements and Rayleigh fading channel for long-haul communications. With above assumptions, either 2×1 or 2×2 virtual MIMO systems are formed for long-haul communication which can be generalized to $N_s \times N_R$ multiple antenna systems, with N_s sensors per cluster and N_R antenna elements at the DGN. Since number of sensors per cluster is assumed to be $N_s = 2$, the space-time block coding scheme used in this paper is the Alamouti scheme [6]. Receiver side local communications can also be considered so that instead of a DGN, another sensor cluster can receive the sensor data acting as a virtual receiver antenna array [12].

B. Energy consumption computation

The two main components of power consumption along a signal path are power consumed by all power amplifiers (P_{PA}) and power consumed by the circuit blocks (P_C) [12], [21]. The P_{PA} term can be approximated as [12]:

$$P_{PA} = (1+\alpha)\, P_{out} \tag{1}$$

where $\alpha = \xi/\eta - 1$ with η being the drain efficiency of the RF power amplifier and ξ is the peak-to-average ratio (PAR). If M-QAM is used, then $\xi = 3\frac{M-2\sqrt{M}+1}{M-1}$ [12]. The transmit power P_{out} can be calculated as

$$P_{out} = \frac{(4\pi)^2\, d^{\kappa} M_l N_f}{G_t G_r \lambda^2} \bar{E}_b R_b, \tag{2}$$

where d is the transmission distance, κ is the signal attenuation parameter, G_t and G_r are the transmitter and receiver antenna gains respectively, λ is the carrier wavelength, M_l is the link margin, N_f-receiver noise figure, $\bar{E}_b$-average energy per bit required for a given bit-error-rate (BER) specification and R_b is the system bit rate. The receiver noise figure N_f is given by $N_f = \frac{N_r}{N_o}$ where N_r is the power spectral density (PSD) of the total effective noise at the receiver input and N_o is the single-sided thermal noise PSD at the room temperature. We may compute $\overline{E}_b$ using the standard expressions given in [21]. Under an AWGN channel for b (with $(M = 2^b)$) even and ≥ 2,

$$\bar{E}_b = \frac{(M-1)\,N_o}{3b}\left[Q^{-1}\left(\frac{\bar{P}_b\ b}{4\left(1-\frac{1}{\sqrt{2^b}}\right)}\right)\right]^2 . \quad (3)$$

for a fixed average bit error rate-$\bar{P}_b$. When b is odd $\bar{E}_b$ can be approximated by dropping the term $\left(1-\frac{1}{\sqrt{2^b}}\right)$ in the denominator of the argument of inverse Q-function in (3).

For a 2×1 Alamouti MISO system, with the assumptions of a Rayleigh fading channel and BPSK modulation, $\bar{E}_b$ can be computed by inverting,

$$\bar{P}_b = \frac{1}{4}\left(1-\frac{1}{\sqrt{1+\frac{1}{\bar{E}_b/2N_o}}}\right)^2\left(2+\frac{1}{\sqrt{1+\frac{1}{\bar{E}_b/2N_o}}}\right) \quad (4)$$

for a fixed $\bar{P}_b$. The bit error rate of an M-ary QAM STBC based $2 \times N_R$ MIMO system with a square constellation (i.e. b is even) is given by, for $b \geq 2$, [21]

$$\bar{P}_b = \frac{4}{b}\left(1-\frac{1}{2^{b/2}}\right)\frac{1}{2^{2N_R}}\left(1-\frac{1}{\sqrt{1+\frac{1}{\bar{E}_b/2N_o}}}\right)^{2N_R} \times \quad (5)$$
$$\sum_{k=0}^{2N_R-1}\frac{1}{2^k}\binom{2N_R-1+k}{k}\left(1+\frac{1}{\sqrt{1+\frac{1}{\bar{E}_b/2N_o}}}\right)^k .$$

When $b \geq 2$ is odd, we will use (5) after dropping the term $\left(1-\frac{1}{2^{b/2}}\right)$ as an upper-bound for the BER. Thus, again by inverting the above equation $\bar{E}_b$ can be computed for a fixed $\bar{P}_b$.

The power consumption in all the circuit blocks of a sensor to broadcast the data during local communications within the cluster can be computed as

$$P_C^{(Loc_{Tx})} \approx P_{DAC} + P_{mix} + P_{filt} + P_{synth} \quad (6)$$

where P_{DAC}, P_{mix}, P_{filt}, P_{synth}, are the power consumption values for the D/A converter (DAC), the mixer, the active filters at the transmitter side and the frequency synthesizer, respectively. Similarly, the circuit power consumption during the reception of the broadcasted data from other sensors of the cluster is

$$P_C^{(Loc_{Rx})} \approx P_{synth} + P_{LNA} + P_{mix} + P_{IFA} + P_{filr} + P_{ADC} \quad (7)$$

where P_{LNA}, P_{IFA}, P_{filr} and P_{ADC} are the powers consumed by Low Noise Amplifier (LNA), the intermediate frequency amplifier (IFA), the active filters at the receiver side and the A/D converter (ADC), respectively.

The total energy consumption per bit for local communication can then be estimated as:

$$E_t^{Local} = \frac{{P_{PA}}^{(Loc)} + P_C^{(Loc_{Tx})} + (N_s-1) \times P_C^{(Loc_{Rx})}}{{R_b}^{(Loc)}} \quad (8)$$

where ${R_b}^{(Loc)}$ is the bit rate for local communication. In this case, ${P_{PA}}^{(Loc)}$ is computed from (1) where $\bar{E}_b$ is computed using (3) since the local channel is assumed to be AWGN and with $d = d_{Loc}$. Note that, d_{Loc} is the distance between any pair of sensors in a cluster.

As DGN is not energy constrained the total circuit energy consumption of a cluster during long-haul communication can be given as

$$P_C^{(Long)} \approx (P_{DAC} + P_{mix} + P_{filt} + P_{synth}) \times N_s \quad (9)$$

where N_s is the number of sensor nodes in a cluster.

For long-haul communication, the total energy consumption per bit for a cluster is approximated as

$$E_t^{Long} = \frac{{P_{PA}}^{(Long)} + P_C^{(Long)}}{{R_b}^{(Long)}} . \quad (10)$$

where ${R_b}^{(Long)}$ is the bit rate for long-haul communication. Again here ${P_{PA}}^{(Long)}$ is computed from (1) with $d = d_{Long}$, where d_{Long} is the distance between the cluster and the DGN. Since the long-haul communication channel is assumed to be Rayleigh, then $\bar{E}_b$ can be computed using either (4) or (5). We assume that the DGN has perfect knowledge of the channel coefficients.

C. System Training for Distributed Compression

Before starting real-time data compression the DGN needs to initialize the existing correlation structure among sensor data so that it can predict future sensor readings. Thus, during the initial training period, N_{tr} uncoded sensor data samples from each sensor are collected by the DGN. When the system is under training the DGN computes the prediction, or side information, of sensor j as a linear combination of the past readings of sensor j along with the current readings of already decoded sensors available at the DGN as follows (Note that, here we have assumed that sensors are ordered according to the order they are decoded, starting with the reference sensor). If $X_k^{(j)}$ is the reading of the sensor j at time k, then the prediction for $X_k^{(j)}$ can be computed at DGN as:

$$Y_k^{(j)} = \sum_{l=1}^{j-1} \alpha_l^{(j)} X_k^{(l)} + \sum_{i=1}^{M} \beta_i^{(j)} X_{k-i}^{(j)} = \underline{\theta}^{(j)T}(k) \underline{z}^{(j)}(k) \tag{11}$$

where $\alpha_l^{(j)}$'s and $\beta_i^{(j)}$'s are the weighting coefficients and we have defined,

$$\begin{aligned} \underline{\theta}^{(j)} &= \left(\alpha_1^{(j)} \; \alpha_2^{(j)} \cdots \alpha_{j-1}^{(j)} \; \beta_1^{(j)} \; \beta_2^{(j)} \; \cdots \; \beta_M^{(j)} \right)^T, \\ \underline{z}^{(j)}(k) &= \left(X_k^{(1)} \; X_k^{(2)} \; \cdots \; X_k^{(j-1)} \; X_{k-1}^{(j)} \; X_{k-2}^{(j)} \; \cdots \; X_{k-M}^{(j)} \right)^T. \end{aligned}$$

(Note that, in numerical results, we considered the side information as a linear combination of the past readings of sensor j and current reading of only the reference sensor since each cluster has only two sensors). The prediction error is

$$e_k^{(j)} = X_k^{(j)} - Y_k^{(j)} = X_k^{(j)} - \underline{\theta}^{(j)}(k)^T \underline{z}^{(j)}(k). \tag{12}$$

In this paper, we propose to update $\underline{\theta}^{(j)}(k)$ for each k using the following RLS algorithm:

$$\underline{\theta}_{new}^{(j)} = \underline{\theta}^{(j)}(k+1) = \underline{\theta}^{(j)}(k) + \underline{g}^{(j)}(k+1) \; e_{k+1}^{(j)}, \tag{13}$$

where $\underline{g}^{(j)}(k+1)$ is the gain vector defined as

$$\underline{g}^{(j)}(k+1) = \frac{\lambda^{-1} \; (R^{(j)}(k))^{-1} \; \underline{z}^{(j)}(k+1)}{1 + \lambda^{-1} \; \underline{z}^{(j)T}(k+1) \; (R^{(j)}(k))^{-1} \; \underline{z}^{(j)}(k+1)} \tag{14}$$

with λ being the exponential weighting factor of the RLS algorithm. Note that, although RLS algorithm increases computational load at the DGN, compared to previously proposed LMS updating, since DGN is assumed to have no energy constraints we believe this is justifiable [20].

During initial system training all sensors share their data with other sensors inside a cluster via local communications. Each sensor combines the data of all the samples for N_{tr} period and then broadcasts the combined information during local communication. Once, each sensor has N_{tr} samples from all the sensors in the cluster it acts as an antenna element of the so formed multiple antenna system. Then STBC is implemented for transmitting data from the cluster to the DGN. Thus, the total energy required for training by all N_s sensor nodes of a cluster is

$$E^{Tr} = (E_t^{Local} + E_t^{Long}) \times n \times N_s \times N_{tr} \tag{15}$$

where n is the bits per sample without compression. At the end of the training period, the DGN initializes the prediction error variance as:

$$\sigma_{e_k^{(j)}}^2 = \frac{1}{N_{tr} - 1} \sum_{k=1}^{N_{tr}} |e_k^{(j)}|^2. \tag{16}$$

D. Correlation Tracking

The distributed compression algorithm requires an underlying codebook which supports multiple compression rates and common to both DGN and each of the sensor as in [18]. Using this codebook a sensor need to use only $i(k,j)$ bits (where $i(k,j)$ is the encoding information from DGN to sensor j at time k) to represent its reading instead of the n bits produced by an n-bit A/D converter where $i(k,j) < n$. The decoding errors will not occur if the prediction $Y_k^{(j)}$ and the actual reading $X_k^{(j)}$ are no further than $2^{i(k,j)-1}\Delta$ apart where Δ is the quantization step of A/D converter.

Thus, if the actual reading and the prediction are not more than $2^{i(k,j)-1}\Delta$ apart, then $X_k^{(j)}$ will be decoded correctly. Using the Chebyshev's inequality the probability of decoding error can be bounded as

$$P[| \; e_k^{(j)} \; | > 2^{i(k,j)-1}\Delta] \leq \frac{\sigma_{e_k^{(j)}}^2}{\left(2^{i(k,j)-1}\Delta\right)^2},$$

where the prediction error $e_k^{(j)}$ is assumed to have a distribution with zero mean and variance $\sigma_{e_k^{(j)}}^2$. We can set the probability of decoding error to be less than a given value

$$P_e = \frac{\sigma_{e_k^{(j)}}^2}{\left(2^{i(k,j)-1}\Delta\right)^2}.$$

Now the required encoding information can be computed as

$$i(k,j) \;\; = \;\; \frac{1}{2} \log_2 \left(\frac{\sigma_{e_k^{(j)}}^2}{\Delta^2 P_e} \right) + 1. \tag{17}$$

At each time instant k, a filtered estimate of $\sigma_{e_k^{(j)}}^2$ can be obtained as

$$\sigma_{e_k^{(j)}}^2 = (1-\gamma)\sigma_{e_{k-1}^{(j)}}^2 + \gamma(e_k^{(j)})^2, \tag{18}$$

where $0 \leq \gamma \leq 1$ is a *forgetting factor* [23].

The $i(k,j)$-bit stream transmitted by a sensor represents the information needed by the DGN in order to decode the actual sensor reading with the help of codebook [18]. At each sample instant k the DGN broadcasts encoding information $i(k,j)$ of all sensor nodes of a cluster. Since $i(k,j) \leq n$, this requires using a maximum of $N_s \times \log_2(n)$ bits. If L is the total number of samples collected from each sensor of a cluster and r is the number of times a cluster receives encoding information updates from the DGN during correlation tracking, then we have that,

$$L = N_{tr} + r \times N_{Rec}. \tag{19}$$

Thus,

$$r = \frac{L - N_{tr}}{N_{Rec}}. \tag{20}$$

The overall energy consumption from all sensors of a cluster in receiving this encoding information can be computed as (since we are only concerned by the energy spent by the data collection sensors),

$$E^{Re} = \frac{r \times P_c^{Rec} \times \log_2(n) \times (N_s^2)}{R_b^{Long}}. \tag{21}$$

where,

$$P_C^{(Rec)} \approx P_{synth} + P_{LNA} + P_{mix} + P_{IFA} + P_{filr} + P_{ADC}. \tag{22}$$

We assume that local and long-haul communications are also performed r times, once in every T_{Rec} time period during correlation tracking. The total energy consumption for transmission of compressed readings by all the sensors of a cluster to the DGN is then given by

$$\begin{aligned} E^{Comp} &= \left(E_t^{Local} + E_t^{Long}\right) \sum_{j=1}^{N_s} \sum_{p=0}^{r-1} \sum_{k=N_{Rec}p+N_{tr}+1}^{N_{Rec}(p+1)+N_{tr}} i(k,j) \\ &= \left(E_t^{Local} + E_t^{Long}\right) \sum_{j=1}^{N_s} \sum_{k=N_{tr}+1}^{L} i(k,j). \end{aligned} \tag{23}$$

As the sensors combine the data of all the samples in each T_{Rec} period before long-haul communication the number of don't care bits need to be added for M-QAM modulation at each transmitting antenna element reduces resulting in increased energy savings. The number of don't care bits required by a sensor for local communication during correlation tracking can be computed as

$$N_{DC}^{local} = \sum_{j=1}^{N_s} \sum_{p=0}^{r-1} \left[N_{Rec}\ i(p\ N_{Rec} + N_{tr} + 1, j)\right] \mod \log_2 M \tag{24}$$

Note that extra energy needs to be spent for transmitting these don't care bits. Similarly, N_{DC}^{long} for long-haul communications can be computed as:

$$N_{DC}^{long} = \sum_{p=0}^{r-1} \left(\left[N_{Rec} \sum_{j=1}^{N_s} i(p\ N_{Rec} + N_{tr} + 1, j) \right] \mod \log_2 M \right). \tag{25}$$

Therefore, the total extra energy spent on don't care bits is,

$$E^{Extra} = E_t^{Local} \times N_{DC}^{local} + E_t^{Long} \times N_{DC}^{long}. \tag{26}$$

E. Energy Efficiency of the Complete System

The total energy consumed by the proposed V-MIMO-based system in collecting all L samples from all the sensors of a cluster can be computed as

$$E^{Total} = E^{Tr} + E^{Re} + E^{Comp} + E^{Extra}. \tag{27}$$

To compute the energy efficiency of the proposed system with respect to three different reference systems. First, is a system where all sensors send their readings to the DGN uncompressed (i.e., SISO with no compression). In this case, the total energy consumption of the system is given by,

$$E^{Ref1} = L \times E_t^{SISO} \times N_s \times (n + n \mod \log_2 M)\,. \tag{28}$$

where E_t^{SISO} is computed using (1) and (6) as

$$E_t^{SISO} = \frac{{P_{PA}}^{(SISO)} + P_C^{(Loc_{Tx})}}{{R_b}^{(SISO)}} \tag{29}$$

where ${R_b}^{(SISO)}$ is the bit rate. In this case ${P_{PA}}^{(SISO)}$ is computed as in (2) with $d = d_{long}$ and [21],

$$\bar{E}_b = \frac{2\,(M-1)\,N_o}{3b} \left(\left(1 - \frac{\bar{P}_b\ b}{2\left(1 - \frac{1}{\sqrt{2^b}}\right)} \right)^{-2} - 1 \right)^{-1} \tag{30}$$

The second reference system is one in which all sensors send their readings using the proposed distributed compression technique but with no virtual transmitter or receiver diversity incorporated (i.e., SISO with compression). In this case, the overall energy expenditure of the system can be given as

$$E^{Ref2} = E^{comp} + E_t^{SISO} \times \left(N_{tr} \times n \times N_s + N_{DC}^{SISO}\right), \tag{31}$$

where

$$E^{Comp} = E_t^{SISO} \times \sum_{k=N_{tr}+1}^{L} \sum_{j=1}^{N_s} i(k,j). \tag{32}$$

The number of don't care bits N_{DC}^{SISO} in this case is computed exactly as in (24). Here we also assume

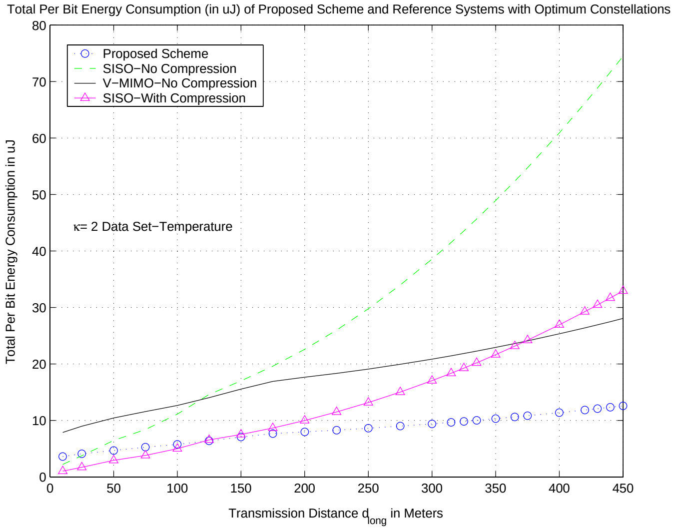

Fig. 4. Total Per Bit Energy Consumption With Optimal Constellations and $\kappa = 2$. (2×1 Virtual STBC)

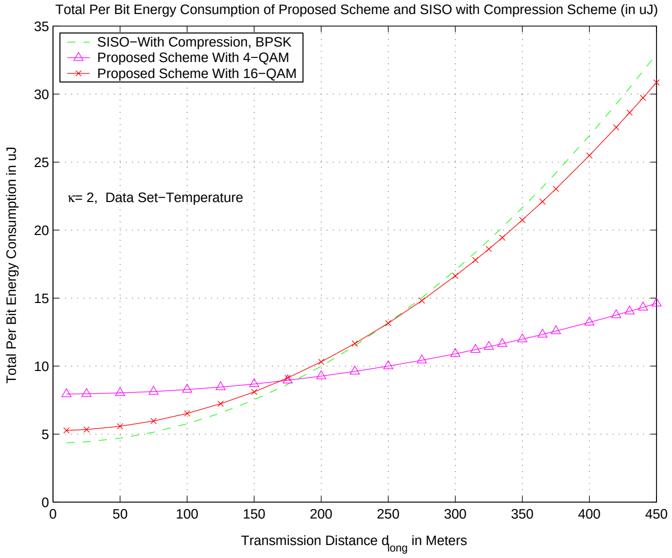

Fig. 5. Per Bit Energy Consumption with $\kappa = 2$ (2×1 Virtual STBC, 4-QAM and 16-QAM for Proposed Scheme and BPSK for SISO)

that sensors combine the N_{Rec} samples after each time sensors receive the encoding information from the DGN.

Finally, we consider a reference system having virtual multiple antenna clusters with no distributed compression (i.e., V-MIMO with no compression). The total energy consumption of this system can be given as

$$E^{Ref3} = (E_t^{Local} + E_t^{Long}) \times N_s \times L \times (n + n \text{ mod } \log_2 M) . \quad (33)$$

We define the energy efficiency η_E of the proposed scheme with respect to any of the above reference systems as

$$\eta_E{}^{Ref} = \frac{E^{Ref} - E^{Total}}{E^{Ref}} \times 100\%. \quad (34)$$

III. EXPERIMENTAL RESULTS

The simulation test bed we considered is a network with 1 DGN and two clusters, each with a pair of equi-distant sensors. One of the 4 sensors always acts as a reference sensor. We fix the local communication distance d_{Loc} (distance among the sensor nodes of any cluster) to 10 meters in all simulations. Simulations are performed for the same sensor data sets of humidity and temperature that were used in [18] over 4000 samples for each sensor. A 12-bit A/D converter with a dynamic range of [-128,128] is assumed for both data sets. All wireless communications are assumed to be performed using M-QAM modulation. AWGN and Rayleigh channels are assumed for local and long-haul communications respectively in all performance simulations. We have chosen $N_{tr} = 65$ and $N_{Rec} = 20$ in all computations. Moreover, a perfect feed back is assumed when the DGN sends the encoding information to the sensor nodes. All power-consumption parameter values are the same as in [12]; i.e., we have assumed $P_{mix} = 30.3$ mW, $P_{filt} = 2.5$ mW, $P_{filr} = 2.5$ mW, $P_{LNA} = 20$ mW, $P_{synth} = 50$ mW, $M_l = 40$ dB, $N_f = 10$ dB, $G_t G_r = 5$ dBi and $\eta = 0.35$.

First, assuming a single antenna element at DGN and two-sensor clusters (2×1 V-MISO) STBC is implemented.

Figure 4 shows the overall energy consumption per information bit of the considered schemes with rate-optimized M-QAM for long-haul communication. Note that, the rate-optimized system pre-computes the optimal constellation size that results in minimum per bit energy consumption for each transmission distance. In Fig. 4, 16-QAM based local transmissions and a very conservative $\kappa = 2$ path loss exponent are assumed [24]. It can be observed from Fig. 4 that the proposed scheme outperforms SISO without compression scheme after ≈ 25 meters, and SISO with compression scheme after ≈ 125 meters. At distances below 125 meters, the energy consumption of the proposed scheme almost matches that of SISO with compression scheme. Hence the proposed scheme with rate optimization can be used for all the transmission distances without considerable loss in energy savings for $d_{Long} < 125$ meters and with significant energy savings for $d_{Long} > 125$ meters.

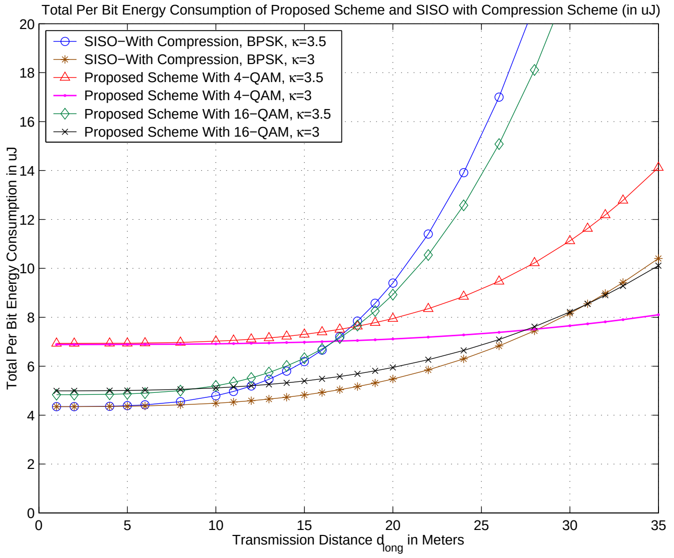

Fig. 6. Per Bit Energy Consumption with $\kappa = 3$ and $\kappa = 3.5$. (2×1 Virtual STBC, 4-QAM and 16-QAM for Proposed Scheme and BPSK for SISO)

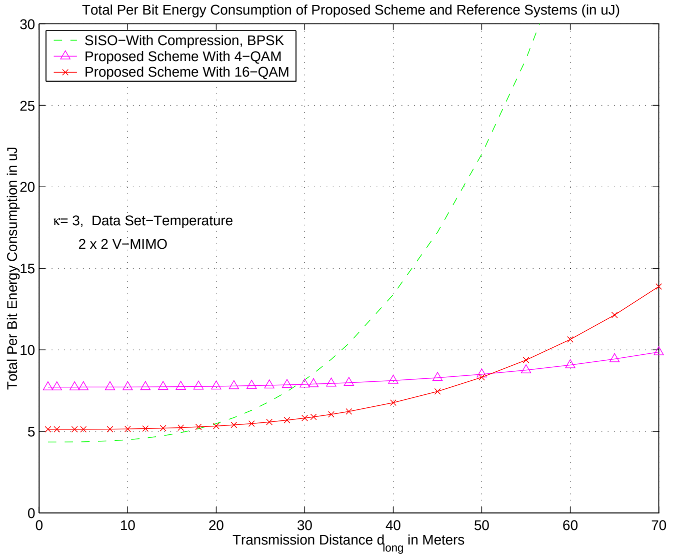

Fig. 7. Per Bit Energy Consumption with $\kappa = 3$. (2×2 Virtual STBC, 4-QAM and 16-QAM for Proposed Scheme and BPSK for SISO)

But optimal variable-rate M-QAM requires complex circuitry at sensor nodes making it unsuitable for low-power wireless sensor networks. In order to overcome this we propose to use sub-optimal, fixed rate M-QAM for both local and long-haul communications. Accordingly Fig. 5 shows the per bit energy consumption for proposed scheme with 4-QAM and 16-QAM for long-haul communications and SISO with BPSK (also using compression). Note that 16-QAM is used for local communication in the proposed scheme. We note that the proposed scheme can be used with 16-QAM for distances below 170 meters, while the 4-QAM modulation can be employed for distances beyond. Such a dual-rate system seems to provide almost the same performance as that of the proposed scheme with optimal variable rates shown in Fig. 4. An alternative for applications requiring the usage of a fixed-rate modulation is to employ 4-QAM for all transmission distances if long-haul communications distance is larger than 170 meters.

However, so far we have been conservative in taking path loss exponent $\kappa = 2$. While $\kappa = 2$ corresponds to free space propagation, near earth propagation encountered in wireless communications may correspond to much larger path loss exponents. In particular usually $3 \leq \kappa \leq 6$ [24]. As we will see below, with such realistic values of κ, the proposed integrated system with V-MIMO outperforms corresponding SISO systems even for very small long-haul communication distance between the DGN and a sensor cluster. Figure 6 illustrates the performance of the above schemes with path loss exponent $\kappa = 3.5$ and $\kappa = 3$, assuming fixed-rate communication. As before again 16-QAM is considered for local communications. Figure 6 shows that proposed scheme with fixed-rate 4-QAM, for example, can provide enormous energy savings over a SISO-based system even for few tens of meters when $\kappa \geq 3$. This makes the proposed virtual MIMO-based wireless sensor network design a good candidate in practice. Figure 7 shows the performance of a 2×2 V-MIMO system with $\kappa = 3$. Figure 7 also assumes fixed rate proposed schemes as before.

Table I shows the energy efficiency of the proposed scheme for a fixed 16-QAM 2×1 V-MIMO system with respect to the reference systems considered. (Note that, ${\eta_E}^1$, ${\eta_E}^2$ and ${\eta_E}^3$ represents the efficiency with respect to SISO-without compression, SISO-with compression and V-MIMO without compression systems, respectively).

IV. CONCLUSIONS

From the above numerical results it can be concluded that the overall energy consumption of the proposed V-MIMO based distributed compression scheme can be much less than that of a corresponding SISO-based system even for transmission distances of few tens of meters in realistic wireless channels. The performance gain of the proposed scheme improves with increasing signal path loss exponent of the channel. Moreover,

TABLE I

ENERGY EFFICIENCIES (IN %) OF THE PROPOSED SCHEME WITH RESPECT TO DIFFERENT REFERENCE SYSTEMS (2×1 VIRTUAL STBC, FIXED-RATE 16-QAM, $\kappa = 3$)

d (m)	10	14	20	24	28	31	35	40	45	50	60	70
η_E^1	2.46	43.67	73.11	81.63	86.34	88.55	90.48	91.96	92.87	93.45	94.12	94.46
η_E^2	-132.02	-35.46	34.81	55.36	66.78	72.13	76.82	80.42	82.62	84.05	85.67	86.51
η_E^3	57.66	57.68	57.73	57.77	57.81	57.84	57.89	57.93	57.97	58.00	58.04	58.06

as we showed with numerical results above, even with fixed-rate systems the proposed V-MIMO based scheme can provide significant performance improvements.

For this work, we have assumed that the DGN has a perfect knowledge of channel coefficients and hence channel estimation is not considered. It is also assumed that the sensors and the DGN are synchronous with each other. In future, we will be considering methods for efficient channel estimation and synchronization suitable for wireless sensor networks of the type considered here.

ACKNOWLEDGMENT

This research was supported in part by Kansas National Science Foundation (NSF) EPSCOR program under the grants KUCR # NSF32223/KAN32224 and KUCR # NSF32195/KAN32196.

REFERENCES

[1] C. S. Raghavendra and S. Singh, "PAMAS - Power aware multi-access protocol with signalling for ad-hoc networks," in *Computer Communications Review*, July 1998.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, pp. 311–335, 1998.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[4] S. K. Jayaweera and H. V. Poor, "Capacity of multi-antenna systems with both receiver and transmitter channel state information," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2697–2709, October 2003.

[5] H. Bolcskei, Arogyaswami, and J. Paulraj, "Multiple-input multiple-output (MIMO) wireless systems," *Chapter in Communications Hand Book of J.Gibson, ed., CRC Press*, pp. 90.1–90.14, April 2002.

[6] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journ. Select. Areas. Commun.*, vol. 16, no. 8, pp. 1451 – 1458, October 1998.

[7] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 5, pp. 1456 – 1467, July 1999.

[8] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 744 – 765, March 1998.

[9] G. J. Foschini, "Layered space-time architecture for wireless comunnication in a flat fading environment when using multi-element antennas," *Bell Labs Tech. Journ.*, vol. 1, no. 2, pp. 41–59, 1996.

[10] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Sig. Proc. Mag.*, vol. 14, pp. 49 – 83, Nov. 1997.

[11] S. Cartreux, P. Driessen, and L. Greenstein, "Simulation results for an interference-limited multple-input multiple-output cellular system," *IEEE Communications Letters*, vol. 4, no. 11, pp. 334–336, November 2000.

[12] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 6, pp. 1089– 1098, October 2004.

[13] M. Dohler, E. Lefranc, and H. Aghvami, "Space-time block codes for Virtual antenna arrays," in *PIMRC, Lisbon, Portugal*, September 2002.

[14] M. Dohler, F. Said, and H. Aghvami, "Higher order space-time block codes for Virtual antenna arrays," in *ICT2003, Tahiti, French Pol.,*, February 2003.

[15] Z. Zeng, M. Dohler, and H. Aghvami, "System performance of a W-CDMA based network with deployed VAA," in *ICT2002, Beijing, China*, June 2002.

[16] M. Dohler, E. Lefranc, and H. Aghvami, "Virtual antenna arrays for future wireless mobile communication systems," in *ICT2002, Beijing, China*, June 2002.

[17] M. Dohler, "Virtual antenna arrays," PhD in Telecommunications, King's College London, London, UK, 2003.

[18] J. Chou, D. Petrovic, and K. Ramchandran, "A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks," in *IEEE INFOCOM*, San Franscisco, CA, Mar. 2003.

[19] M. L. Chebolu, V. K. Veeramachaneni, S. K. Jayaweera, and K. R. Namuduri, "An improved adaptive signal processing approach to reduce energy consumption in sensor networks," in *CISS conference on Information Sciences and Systems*. Princeton, NJ, Mar. 2004.

[20] M. Chebolu and S. K. Jayaweera, "Energy efficiency analysis of an RLS-based adaptive signal processing algorithm for energy aware sensor networks," in *2nd International Conf. on Intelligent Sensing and Information Processing (ICISIP 05)*, Chennai, India, Jan 2005.

[21] S. K. Jayaweera, "Energy efficient Virtual MIMO-based cooperative communications for wireless sensor networks," in *2nd International Conf. on Intelligent Sensing and Information Processing (ICISIP 05)*, Chennai, India, Jan 2005.

[22] ——, "An energy-efficient Virtual MIMO communications architecture based on V-BLAST processing for distributed wireless sensor networks," in *1st IEEE International Conf. on Sensor*

and Ad-hoc Communications and Networks (SECON), Santa Clara, CA, oct 2004.

[23] S. Haykin, *Adaptive Filter Theory*. Upper-saddle River, NJ, USA: Prentice Hall, 1996.

[24] S. K. Jayaweera, “Energy analysis of mimo techniques in wireless sensor networks,” in *38th Annual Conference on Information Sciences and Systems (CISS 04)*, Princeton, NJ, March 2004.