Bayesian Data and Channel Joint Maximum-Likelihood Based Error Correction in Wireless Sensor Networks

Ashish Katiyar
Department of Electrical Engineering
Indian Institute of Technology Rajasthan
Jodhpur, India 342011
ashishkatiyar@iitj.ac.in

Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, India 208016
adityaj@iitk.ac.in

ABSTRACT

We propose a novel Bayesian error correction algorithm based on joint channel and data maximal-likelihood (ML) detection in wireless sensor networks (WSN). The proposed algorithm employs the temporal correlation of the narrowband sensor data in conjunction with the channel state information (CSI) for detection and error correction of the data received over the Rayleigh fading wireless channel. The proposed joint maximum-likelihood (JML) algorithm compares the joint channel and data likelihoods along different paths of the data likelihood tree (DLT), which is readily adaptable for efficient practical implementation in WSNs. Further, the JML scheme employs the sphere decoder for computation of the maximally likely sphere sensor data vectors in the WSN and thus has a low computational complexity. Simulation results demonstrate significantly reduced sensor error for the proposed WSN sensor correction technique over competing schemes existing in current literature.

Categories and Subject Descriptors

C.2 [**Computer-Communication Networks**]: Wireless Communication; C.4 [**Performance of systems**]: Fault tolerance

General Terms

Reliability, Performance

Keywords

Wireless Sensor Network (WSN), Likelihood, Sphere Decoder

1. INTRODUCTION

Wireless Sensor Networks (WSN) have attracted the attention of researchers in diverse fields of interest due to their key role in a wide variety of applications such as habitat and environment monitoring, military defense, disaster warning systems, agriculture [10, 7] etc. These applications involve the sensing of an appropriate phenomenon such as temperature, pressure, radiation etc. over a vast region and transmitting the sensor measurements to a cluster head for processing and mining of the received raw information. Hence, a WSN consists of a large number of networked sensor nodes transmitting the sensor data to the cluster head or fusion center over wireless links. During this process of transmission and reception, errors arise in received data due to noise at the receiver and the fading nature of the wireless channel [9].

Data reliability is a critically important factor for the practical implementation of WSNs, especially in automation and hazard detection applications. A practical constraint in the deployment process of a WSN is the cost of the sensor node, since WSNs typically necessitate a large number of networked sensor nodes. In such scenarios, using high quality sensor nodes, with built in hardware/software capability for reliability, results in a prohibitively substantial increase in the total cost of deployment, making it practically infeasible. Hence, robust error correction algorithms at the fusion center are the key to ensuring reliability of the received sensor information in a WSN. They enable the employment of low cost sensor nodes, effectively relegating the processing complexity to the few cluster heads, thus ensuring low cost and scalability of the WSN.

In this context, we propose a robust sensor data rectification algorithm towards enhancing the reliability of the WSN sensor data detection. The proposed algorithm harnesses the inherent temporal correlation of the narrowband phenomenon being sensed in conjunction with the channel state information available at the wireless receiver for effective error correction. Our scheme is based on efficiently computing the joint model and channel maximum-likelihood (JML) estimate of the received sensor data. Further, we employ the sphere-decoder [1] for fast computation of the maximally likely sensor data vectors. The multi-sample likelihood is computed employing a novel data likelihood tree (DLT) for fast implementation, resulting in a significant reduction in the computational complexity. The joint ML estimate of the sensor data is computed by choosing the DLT path corresponding to the lowest likelihood cost metric.

The advantage of the proposed JML approach compared to other competing schemes for data correction in WSNs, such as the prediction history tree (PHT) based scheme proposed in [6], is that it does not critically depend on the model accuracy. This is important since estimation of highly accurate models in WSNs requires a large number of training samples with high transmission power, adding to the power and data overhead and results in longer delays and lower battery life. Moreover, such a highly accurate model is unnecessary, since one can intelligently employ it in conjunction with the channel state information to design a robust and highly accurate overall error correction algorithm. Also, it can be seen from the simulation results that the proposed algorithm yields superior error correction performance even for comparatively wideband i.e. lower temporal correlation sensor data samples and is therefore not significantly sensitive to the accuracy of the Bayesian prior data model.

The presentation in the rest of the paper is organized as follows. Section 2 presents the WSN system model. Section 3 details the proposed JML algorithm for error correction in WSNs along with the DLT structure for likelihood comparison and the modified sphere-decoder algorithm, which aids in fast likelihood computation of the maximally likely sensor data vectors. Simulation results and performance comparison with competing error correction schemes are given in section 4 and we conclude with section 5.

2. WSN SYSTEM MODEL

The sensed data at the nodes of the WSN is transmitted to the cluster head over the fading wireless channel. Let $\theta(k)$ denote the k^{th} sensed sample, which is mapped to a transmit symbol vector $\mathbf{x}(k)$ by the encoding function $f : \mathbb{R} \to \mathbb{C}^m$ as,

$$\mathbf{x}(k) = [x_1(k), x_2(k), \ldots, x_m(k)]^T \triangleq f(\theta(k)),$$

where the modulated symbols $x_i(k)$, $1 \le k \le m$ are drawn from the source constellation $\mathcal{S}$ consisting of symbols $s_j, 1 \le j \le N_s$. For instance, for a source employing QPSK modulation, the symbol set $\mathcal{S}$ given as,

$$\mathcal{S} = \left\{ \sqrt{\frac{P_d}{2}} \left(u_I + ju_Q\right) \middle| \, u_I, u_Q \in \{-1, +1\} \right\},$$

where P_d is the power of the data symbols. The received data $y_i(k)$ at the cluster head at time instant k can be represented as,

$$y_i(k) = hx_i(k) + n(k),$$

where $x_i(k)$ is the i^{th} transmitted modulated symbol corresponding to the sensor data sample at the k^{th} sampling instant, h_i is the complex baseband fading channel coefficient corresponding to the wireless channel between the sensor and cluster head and $n(k)$ is the noise at the receiver. Employing the output symbol $y_i(k)$ and the channel state information h, the optimal maximum-likelihood scheme for detection of the transmitted symbol $x_i(k)$ is given as,

$$\hat{x}_i(k) = \arg\max_{s \in \mathcal{S}} \|y_i(k) - hs\|^2, \tag{1}$$

where $\hat{x}_i(k)$ is the detected symbol corresponding to the transmitted symbol $x_i(k)$ [8] [2]. The decoded sensor data $\hat{\theta}(k)$ corresponding to the detected symbols $\hat{x}_i(k)$ is given by the decoder function f^{-1} as,

$$\hat{\theta}(k) = f^{-1}(\hat{x}_1(k), \hat{x}_2(k), \ldots, \hat{x}_m(k)) = f^{-1}(\hat{\mathbf{x}}(k)), \tag{2}$$

where $\hat{\mathbf{x}}(k) = [\hat{x}_1(k), \hat{x}_2(k), \ldots, \hat{x}_m(k)]^T$ is the detected symbol vector. The channel noise $n(k)$ and the deep fade of the wireless channel h can result in the erroneous detection of the transmitted symbols $x_i(k)$, introducing errors in the reported sensor data. A major disadvantage of the above detection scheme is that it does not consider the high level of spatio-temporal correlation inherent in the narrowband sensing phenomenon. Hence, motivated by this observation, we propose the Bayesian joint model and channel maximum-likelihood (JML) scheme below to exploit prior statistical information towards sensor data error minimization. Before we proceed, it is worth noting that the decoupled detection problem in (1) can be represented by the equivalent multi-dimensional detection problem,

$$\hat{\mathbf{x}}(k) = \arg\max_{\mathbf{s} \in \mathcal{S}^m} \|\mathbf{y}(k) - \mathbf{H}\mathbf{s}\|^2,$$

where $\mathbf{y}(k) = [y_1(k), y_2(k), \ldots, y_m(k)]^T$ is the received symbol vector, and $\mathcal{S}^m$ denotes the multi-dimensional transmit symbol constellation $\mathcal{S}^m \triangleq \mathcal{S} \times \mathcal{S} \times \ldots \times \mathcal{S}$. The channel matrix $\mathbf{H}$ is the $m \times m$ diagonal matrix of the fading channel coefficient h defined as,

$$\mathbf{H} = \begin{bmatrix} h & 0 & \cdots & 0 \\ 0 & h & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & \cdots & h \end{bmatrix}.$$

2.1 Bayesian WSN Error Correction

We begin by describing in detail the data prediction model and the various factors affecting the accuracy of the model. It is well known in literature that the a narrowband time-series $\theta(k)$ can be predicted by the optimal auto-regressive (AR) Wiener filter [3] given as,

$$\begin{aligned} \theta(k) &= a_1\theta(k-1) + a_2\theta(k-2) + \ldots + \theta_L y(k-L) + w(k), \\ &= \mathbf{a}^T \bar{\theta}(k-1), \end{aligned} \tag{3}$$

where w is the modeling error with variance $\sigma_w^2 = \mathrm{E}\left\{|w|^2\right\}$. The optimal modeling mean-square error (MSE) minimizing AR coefficient vector $\mathbf{a} \triangleq [a_1, a_2, ..., a_L]^T \in \mathbb{R}^L$ can be readily obtained as the solution of the Wiener-Hopf equations [4], [5]. Hence, such an optimal data model can be conveniently employed to harness the temporal correlation existing in the sensor data. An AR model of very high accuracy can be constructed using a very low number of training samples. Further, since the optimal coefficients are computed offline and depend only on the second-order statistics of the sensor data, this model has a very low computational complexity. Moreover, the accuracy of the model can be readily adapted to the level of temporal correlation and complexity limits by varying the order L of the prediction model above.

Let $\mathbf{X}_1^{N_b} = [\mathbf{x}(1), \mathbf{x}(2), \ldots, \mathbf{x}(N_b)] \in \mathbb{C}^{m \times N_b}$ and $\mathbf{Y}_1^{N_b} = [\mathbf{y}(1), \mathbf{y}(2), \ldots, \mathbf{y}(N_b)] \in \mathbb{C}^{m \times N_b}$ denote the transmitted and received symbol blocks respectively, where N_b denotes the block length. Hence, employing the prediction model above, the expression for the joint Bayesian data and channel likelihood of $\mathbf{X}_1^{N_b}, \mathbf{Y}_1^{N_b}$ is given by the Gaussian likelihood function,

$$\mathcal{L}\left(\mathbf{Y}_1^{N_b}|\mathbf{H};\mathbf{S}_1^{N_b}\right) = \prod_{i=1}^{N_b} e^{-\frac{1}{\sigma_n^2}\|\mathbf{y}(i)-\mathbf{H}\mathbf{s}(i)\|^2} \times \prod_{L+1}^{N_b} e^{-\frac{1}{\sigma_w^2}\left\|f^{-1}(\mathbf{s}(i))-\mathbf{a}^T\mathbf{F}^{-1}\left(\mathbf{S}_{i-L}^{i-1}\right)\right\|^2} \quad (4)$$

where $\mathbf{S}_{i-L}^{i-1}$ denotes the block $\left[\mathbf{s}\left(i-L\right), \mathbf{s}\left(i-L+1\right), \ldots, \mathbf{s}\left(i-1\right)\right]$ and $\mathbf{F}^{-1}: \mathbb{C}^{m\times L} \rightarrow \mathbb{R}^L$ denotes the joint decoder function defined as,

$$\mathbf{F}^{-1}\left(\mathbf{S}_{i-L}^{i-1}\right) \triangleq \left[f^{-1}\left(\mathbf{s}_{i-1}\right), f^{-1}\left(\mathbf{s}_{i-2}\right), \ldots, f^{-1}\left(\mathbf{s}_{i-L}\right)\right]^T.$$

In the total likelihood expression given in (4) above, the first term represents the channel likelihood and second term represents the prior Bayesian data likelihood of the encoded sensor sample block $\mathbf{S}_{i-L}^{i-1}$, which corresponds to the temporal narrowband nature of the sensor data samples. The corresponding negative log-likelihood cost metric for the above likelihood expression, which is more amenable to block detection, can be derived as,

$$\mathcal{G}\left(\mathbf{Y}_1^{N_b}|\mathbf{H};\mathbf{S}_1^{N_b}\right) = \frac{1}{\sigma_n^2}\sum_{i=1}^{N_b}\|\mathbf{y}\left(i\right)-\mathbf{H}\mathbf{s}\left(i\right)\|^2 + \frac{1}{\sigma_w^2}\sum_{L+1}^{N_b}\left\|f^{-1}\left(\mathbf{s}_i\right)-\mathbf{a}^T\mathbf{F}^{-1}\left(\mathbf{S}_{i-L}^{i-1}\right)\right\|^2.$$

The optimal Bayesian maximum-likelihood estimate $\hat{\mathbf{X}}_1^{N_b}$ is given as the solution of the above likelihood cost minimizer, which can be stated formally as,

$$\hat{\mathbf{X}}_1^{N_b} = \underset{\mathbf{S}_1^{N_b}\in(\mathcal{S}^m)^{N_b}}{\arg\min}\ \mathcal{G}\left(\mathbf{Y}_1^{N_b}|\mathbf{H};\mathbf{S}_1^{N_b}\right) \quad (5)$$

In principle, finding the optimal solution of the above block coupled likelihood cost-function yields the ML decoded block $\hat{\mathbf{X}}_1^{N_b}$, from which the sensor sample estimates $\hat{\theta}\left(k\right)$ can be computed as $\hat{\theta}\left(k\right) = f^{-1}\left(\hat{\mathbf{x}}\left(k\right)\right)$. However, the dimension of the state space $\left(\mathcal{S}^m\right)^{N_b}$ grows exponentially in mN_b, where m is the resolution of the encoder $f\left(\cdot\right)$ and N_b is the block length, thus rendering the problem NP-hard. For instance, consider a 16-bit encoder (i.e. 8 symbol QPSK modulator), with a block length $N_b = 100$. The dimension of the complete state space is $2^{16\times 100}$, which is practically impossible to optimize. Hence, we present a novel low-complexity scheme based on the data likelihood tree (DLT) structure in the section below, which computes the optimal Bayesian JML estimate corresponding to the cost function in (5).

3. JML BASED ERROR CORRECTION

Consider the DLT structure for Bayesian ML detection given in Fig.1. The DLT shown therein comprises of $N_T = 3$ levels, labeled as $l = 0, 1, 2$. At the current time instant n, let the root node at level $l = 0$ denote the latest corrected symbol vector $\hat{\mathbf{x}}_c\left(n-2\right)$ as the output of the optimal JML detector. At this stage, the DLT is employed to compute the likelihood cost corresponding to the state subspace at time instant $n-1$, followed by choosing the corrected transmit symbol vector $\hat{\mathbf{x}}_c\left(n-1\right)$. This state subspace is of dimension $|\mathcal{S}|^m$. A significant number of the states can be discarded probabilistically due to the low-likelihood state metric, derived from the likelihood computation procedure. However,

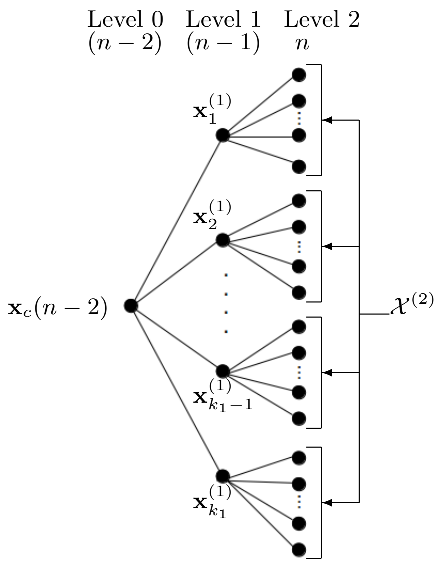

Figure 1: Bayesin Data Likelihood Tree (DLT)

direct computation of the likelihood metric at every state results in a drastic increase in the computational complexity, due to the large cardinality of the state subspace. In such a scenario, the sphere-decoding algorithm [1] can be readily employed to compute the set $\mathcal{X}^{(1)}$ of k_1 of maximally-likely vectors at $l = 1$ given as $\mathcal{X}^{(1)} = \left\{\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \ldots, \mathbf{x}_{k_1}^{(1)}\right\}$, which satisfy the channel cost-metric condition,

$$\mathcal{X}^{(1)} = \left\{\mathbf{s}\in\mathcal{S}^m|\ \|\mathbf{y}(n-1)-\mathbf{H}\mathbf{s}\|^2 \leq d^2\right\}, \quad (6)$$

where d is the sphere radius. Thus, level $l = 1$ in essence consists of the all the vectors with significant likelihood, corresponding to the k_1 nodes at $l = 1$ in the DLT, obtained by passing the observed vector $\mathbf{y}\left(n-1\right)$ through the sphere decoder. Level $l = 2$ in turn contains the set of vectors obtained by passing the observed vector $\mathbf{y}\left(n\right)$ at time instant n through the sphere decoder. This set of size k_2 is represented as $\mathcal{X}^{(2)} = \left\{\mathbf{x}_1^{(2)}, \mathbf{x}_2^{(2)}, \ldots, \mathbf{x}_{k_2}^{(2)}\right\}$ and corresponds to the sphere-vectors $\|\mathbf{y}(n)-\mathbf{H}\mathbf{s}\|^2 \leq d^2$. The DLT computes and populates the total Bayesian likelihood entries corresponding to the k_1k_2 paths from the level 0 root note to each node at level 2 corresponding to time instant n, which is the sum of the log-likelihoods of the individual paths from level 0 to level 1 and from level 1 to level 2. The JML detected symbol $\hat{\mathbf{x}}_c\left(n-1\right)\in\mathcal{X}^{(1)}$ for time instant $n-1$ can now be computed with relatively low computation complexity as given in (7).

Thus, the DLT based Bayesian JML scheme effectively chooses corrected symbol vector $\hat{\mathbf{x}}_c\left(n-1\right)$ corresponding to the sample at time $n-1$ through a fast low-complexity likelihood computation scheme. Subsequently, the node corresponding to $\hat{\mathbf{x}}_c\left(n-1\right)$ is set as the root node at $l = 0$, with the nodes at subsequent levels $l = 1, 2$ corresponding to the maximally-likely sphere vectors of time instants $n, n+1$. This involves population of the $l = 2$ nodes with the appropriate likelihood values through computation of the sphere-vectors corresponding to the time instant $n+1$. This procedure is repeated until the entire block of N_b sensor symbol vectors and the corresponding sensor samples are decoded. The

$$\hat{\mathbf{x}}_c(n-1) = \mathbf{s}^{(1)} = \underset{\substack{\mathbf{s}^{(1)} \in \mathcal{X}^{(1)} \\ \mathbf{s}^{(2)} \in \mathcal{X}^{(2)}}}{\arg\min} \left\{ \frac{1}{\sigma_n^2} \sum_{i=1}^{2} \left\| \mathbf{y}(n+i-2) - \mathbf{H}\mathbf{s}^{(i)} \right\|^2 + \frac{1}{\sigma_w^2} \sum_{i=1}^{2} \left\| f^{-1}\left(\mathbf{s}^{(i)}\right) - \mathbf{a}^T \mathbf{F}^{-1}\left(\bar{\mathbf{G}}^{(i)}\right) \right\|^2 \right\}, \tag{7}$$

$$\bar{\mathbf{G}}^{(1)} = \begin{bmatrix} f^{-1}\left(\hat{\mathbf{x}}_c(n-2)\right) \\ f^{-1}\left(\hat{\mathbf{x}}_c(n-3)\right) \\ \vdots \\ f^{-1}\left(\hat{\mathbf{x}}_c(n-L-1)\right) \end{bmatrix}, \bar{\mathbf{G}}^{(2)} = \begin{bmatrix} f^{-1}\left(\mathbf{s}^{(1)}\right) \\ f^{-1}\left(\hat{\mathbf{x}}_c(n-2)\right) \\ \vdots \\ f^{-1}\left(\hat{\mathbf{x}}_c(n-L)\right) \end{bmatrix}.$$

sensor sample corresponding to each corrected sensor symbol vector is given by the decoder in (2). To complete the discussion, we give the sphere-decoding algorithm [1] in this context of Bayesian sensor data correction for WSNs.

3.1 Sphere-Decoder for Sensor Data Likelihood Computation

The sphere-decoder is a highly effective tool to compute the set of transmit vectors lying within a sphere of radius d corresponding to a received vector $\mathbf{y}$, in other words, corresponding to a given log-likelihood threshold. As given in (6), it can be employed to compute the maximally likely sphere vectors for the received vector $\mathbf{y}$ at each level l of the DLT. Hence, it can be recast as a search for all vectors $\mathbf{s} \in \mathcal{S}^m$, satisfying the constraint,

$$\|\mathbf{y} - \mathbf{H}\mathbf{s}\|^2 < d^2.$$

Since the matrices $\mathbf{y}, \mathbf{H}, \mathbf{s}$ contain complex entries in a typical baseband wireless communication system model representation, we recast the matrices as the following stacked matrices of real entries, rendering them amenable to the implementation of the sphere decoding algorithm. Let the quantities above be expressed in terms of real and imaginary components as, $\mathbf{y} = \mathbf{y}_r + j\mathbf{y}_i$, $\mathbf{H} = \mathbf{H}_r + j\mathbf{H}_i$, $\mathbf{s} = \mathbf{s}_r + j\mathbf{s}_i$. Hence, the cost metric $\|\mathbf{y} - \mathbf{H}\mathbf{s}\|$ can be rewritten as,

$$\begin{aligned} \|\mathbf{y} - \mathbf{H}\mathbf{s}\| &= \left\|(\mathbf{y}_r + j\mathbf{y}_i) - (\mathbf{H}_r + j\mathbf{H}_i)(\mathbf{s}_r + j\mathbf{s}_i)\right\| \\ &= \left\| \underbrace{\begin{bmatrix} \mathbf{y}_r \\ \mathbf{y}_i \end{bmatrix}}_{\mathbf{y}^t} - \underbrace{\begin{bmatrix} \mathbf{H}_r & -\mathbf{H}_i \\ \mathbf{H}_i & \mathbf{H}_r \end{bmatrix}}_{\mathbf{H}^t} \underbrace{\begin{bmatrix} \mathbf{s}_r \\ \mathbf{s}_i \end{bmatrix}}_{\mathbf{s}^t} \right\|. \end{aligned}$$

In the algorithm below, the individual elements of the vector $\mathbf{s}^t$ are represented as s_i^t. Let the standard $\mathbf{QR}$ decomposition of the matrix $\mathbf{H}^t$ be given as, $\mathbf{H}^t = \mathbf{Q}^t\mathbf{R}^t$. Let the vector $\mathbf{g}$ be defined as $\mathbf{g} \triangleq \left(\mathbf{Q}^t\right)^H \mathbf{y}^t$, and its individual elements be represented as g_i. The elements of the matrix $\mathbf{R}^t$ are represented as $r_{i,j}$. Below, in Algorithm 1, we describe the sphere decoding algorithm adapted to this context of WSN error correction.

Algorithm 1 SPHERE VECTORS

Require: $\mathbf{Q}^t, \mathbf{R}^t, \mathbf{y}^t, d, \mathbf{g} \triangleq \left(\mathbf{Q}^t\right)^H \mathbf{y}^t$;

1: $c = 2m, \tilde{d}_c = d, g_{2m|2m+1} = g_{2m}$;
2: $\mathrm{UB}\left(s_c^t\right) = \min\left(\lfloor \frac{\tilde{d}_c + g_{2m|2m+1}}{r_{c,c}} \rfloor, 1\right)$
3: $s_c^t = -3$;
4: **if** $c > m$ **then**
5: $\quad s_c^t = 0$;
6: **else**
7: $\quad s_c^t = s_c^t + 2$;
8: **end if**
9: **if** $c <= m$ **then**
10: $\quad$ `repeat` = 0;
11: **end if**
12: **if** $\left(\left(s_c^t < \mathrm{UB}\left(s_c^t\right)\right)$ **and** (!`repeat`)$\right)$ **then**
13: $\quad$ GOTO 24;
14: **else**
15: $\quad$ GOTO 17;
16: **end if**
17: $c \leftarrow c + 1$
18: **if** $c = 2m + 1$ **then**
19: $\quad$ TERMINATE
20: **else**
21: $\quad$ `repeat` = 1;
22: $\quad$ GOTO 4;
23: **end if**
24: **if** $c = 1$ **then**
25: $\quad$ GOTO 33;
26: **else**
27: $\quad c = c - 1$;
28: $\quad g_{c|c+1} = g_c - \sum_{j=c+1}^{2m} r_{c,j} s_j^t$;
29: $\quad \tilde{d}_c = \sqrt{\tilde{d}_{c+1}^2 - \left(g_{c+1|c+2} - r_{c+1,c+1} s_{c+1}^t\right)^2}$;
30: **end if**
31: //SOLUTION FOUND;
32: //SAVE $\left(\mathbf{s}^t\right)$ AT LEVEL i AND ITS LIKELIHOOD;
33: $\mathcal{X}^{(i)} \leftarrow \mathcal{X}^{(i)} \cup \left\{\mathbf{s}^t\right\}$;
34: `repeat` = 1;
35: GOTO 4;

4. SIMULATION RESULTS

In this section, we present simulation results and performance comparison of the proposed WSN error correction scheme with competing algorithms. The wireless channel is Rayleigh fading in nature. We present the performance of the proposed scheme for different levels of noise power σ_n^2 and modeling error σ_w^2. We employ a 3 stage model for the generation of the sensor data, corresponding to the time intervals $0 \leq k \leq 300$, $300 \leq k \leq 700$ and $700 \leq k \leq 1000$ respectively, for a total block length of $N_b = 1000$ reported sensor samples at the cluster head. The model parameters for sensor data modeling are shown in table 1. In our simulations, we employ 70 data samples for training the Bayesian data model, i.e. for estimation of the model coefficient vec-

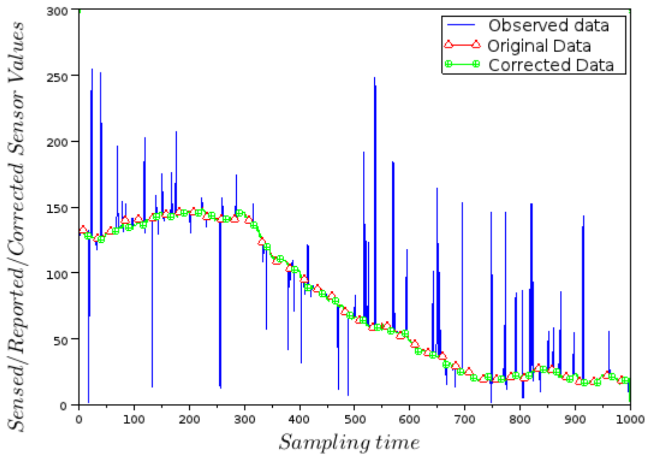

Figure 2: Error Correction performance

Stage	a_1	a_2
First Stage (0-300)	0.7	0.3
Second Stage (300-700)	0.8	0.195
Third Stage (700-1000)	0.75	0.25

Table 1: Model Parameters used for Original data simulation

tor **a** given in (3). The estimated 3 stage generation model parameters are presented in table 2. Fig.2 shows the sensor data and the received erroneous reports at the cluster head. Further, the corrected data following the application of the proposed Bayesian scheme for error correction is also shown therein. It can be readily seen that the corrected data agrees closely with the sensed data, validating the efficacy of the proposed algorithm in the context of WSN. Next we consider the performance of the proposed scheme for varying receiver noise σ_n^2. To characterize the WSN data correction error, we employ the mean-squared error (MSE) metric defined as,

$$\text{MSE} = \frac{1}{N_b} \left\{ \sum_{i=1}^{N_b} \left(\theta_c\left(i\right) - \theta\left(i\right)\right)^2 \right\},$$

where N_b is the total number of samples, $\theta_c\left(i\right)$ is the corrected sensor data subsequent to application of the JML correction algorithm and $\theta\left(i\right)$ is the original sensor value. The performance of the algorithm with varying receiver noise power is shown in Fig.3. Finally, we plot the performance of the proposed algorithm for different levels of modeling error σ_w^2 in the original narrowband sensor data and compare it with that of the existing PHT algorithm [6]. From figure 4

Stage	a_1	a_2
First Stage (0-300)	0.74843	0.251549
Second Stage (300-700)	0.91792	0.0820481
Third Stage (700-1000)	0.707529	0.291461

Table 2: Predicted model parameters

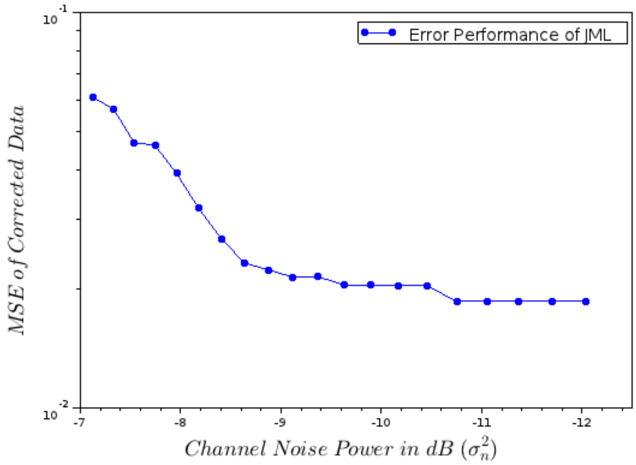

Figure 3: MSE Performance evaluation of error correction for different levels of channel noise σ_n^2.

it can be readily seen that the proposed JML scheme has a much superior MSE performance compared to the existing PHT scheme. This is because, the PHT scheme does not employ the channel likelihood, thereby discarding a significant amount of information that can beneficially aid in the correction of the sensor data. Further, it can be seen that as the modeling error increases, the MSE for PHT algorithm increases, clearly demonstrating the high sensitivity of the PHT scheme to the modeling accuracy. However, the MSE distortion of the corrected data for the proposed JML algorithm remains fairly constant, thereby demonstrating that an approximate prior sensor data model can be successfully employed to drastically reduce the reported data error.

5. CONCLUSION

The proposed JML algorithm in the context of WSNs results in a significant reduction in the error of the reported sensor data. This is achieved by advantageously harnessing the temporal correlation inherent in the narrowband sensor data and employing it along with the channel state information for joint decoding of the received data. Further, the proposed algorithm employs the sphere-decoder for fast computation of the maximially likely sensor data vectors in conjunction with a proposed data likelihood tree (DLT) based low-complexity likelihood computation leading to the choice of the optimal Bayesian corrected estimate of the reported sensor data. Hence, the proposed algorithm has very low complexity and can be readily adapted in practical WSNs for data fusion. Moreover, the performance of the proposed algorithm is not sensitive to the accuracy of the data model, ensuring robustness for practical implementation. Simulation results presented clearly demonstrate the significantly superior performance of the proposed scheme for WSN error correction compared to current schemes in literature such as PHT.

6. REFERENCES

[1] B. Hassibi and H. Vikalo. On the sphere-decoding algorithm i. expected complexity. *Signal Processing,*

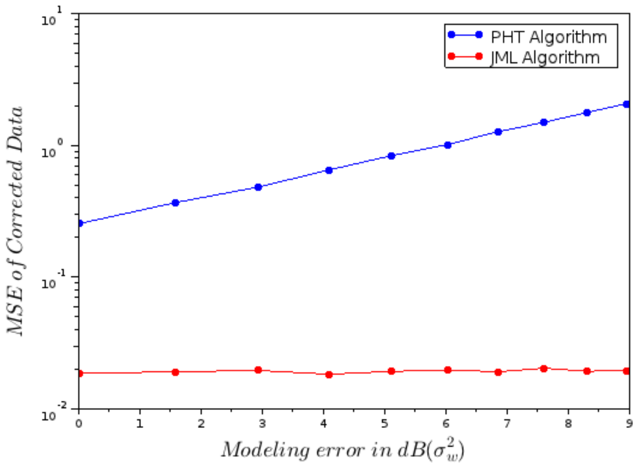

Figure 4: Performance comparison of JML and PHT schemes for varying modeling error σ_w^2.

IEEE Transactions on, 53(8):2806 – 2818, aug. 2005.

[2] S. Kay. *Fundamentals Of Statistical Signal Processing*. Number v. 1. Prentice Hall, 2001.

[3] D. Manolakis, V. Ingle, and S. Kogon. *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*. Artech House signal processing library. Artech House, 2000.

[4] D. Manolakis and J. Proakis. *Digital Signal Processing Principles Algorithms And Applications*. Phi, 2002.

[5] T. Moon and W. Stirling. *Mathematical methods and algorithms for signal processing*. Prentice Hall, 2000.

[6] S. Mukhopadhyay, C. Schurgers, D. Panigrahi, and S. Dey. Model-based techniques for data reliability in wireless sensor networks. *Mobile Computing, IEEE Transactions on*, 8(4):528 –543, april 2009.

[7] A. Swami. *Wireless sensor networks: signal processing and communications perspectives*. J. Wiley, 2007.

[8] H. Trees. *Detection, estimation, and modulation theory*. Wiley, 1968.

[9] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[10] J. Wilson. *Sensor technology handbook*. Number v. 1 in Electronics & Electrical. Elsevier, 2005.