Integrated Source-Channel Decoding for Correlated Data-Gathering Sensor Networks

Sheryl L. Howard
EE Department
Northern Arizona University
Flagstaff, AZ 86001
sheryl.howard@nau.edu

Paul G. Flikkema
EE Department
Northern Arizona University
Flagstaff, AZ 86001
paul.flikkema@nau.edu

***Abstract*— This paper explores integrated source-channel decoding, driven by wireless sensor network applications where correlated information acquired by the network is gathered at a destination node. The collection of coded measurements sent to the destination, called a source-channel product codeword, has redundancy due to both correlation of the measurements and the channel code used for each measurement. At the destination, source-channel (SC) decoding of this code combines decoding using (i) the deterministic structure of the channel-coded individual measurements and (ii) the probabilistic structure of a prior model, called the global model, that describes the correlation structure of the SC product codewords. We demonstrate the utility of SC decoding via MAP SC decoding experiments using a (7,4,3) Hamming code and a Gaussian global model. We also show that SC decoding can exploit even the simplest possible code, a single-parity check code, using a MAP SC decoder that integrates the parity check constraint and global model. We describe the design of a low-complexity message-passing decoder and show it can improve performance in the poor-quality channels often found in multi-hop wireless data-gathering sensor networks.**

I. Introduction

Wireless sensor networks are assemblies of unreliable components: sensors sample their environments with noisy, error-prone transducers, and this information is communicated over unreliable wireless communication links. Moreover, a sensor may fail due to loss of energy or failure of its hardware or software components. Reliability can be improved with increased expenditure of bandwidth, energy, and funds; however, many applications are cost-sensitive, so improving the reliability of practical sensor network systems is of interest.

Consider the data-gathering scenario where sensors sample their environments and collaboratively forward the measurements to one or more destinations or fusion centers via multi-hop wireless networking.[1] The ultimate goal is an estimate with reasonable fidelity along with an estimate of that fidelity. In current implementations, source and channel codes are independently designed based on the separation theorem of information theory. That is, sensors use channel coding to protect transmitted information, and source coding to minimize the cost of transmission. However, in a cost- and energy-constrained network, this may not be optimal:

[1]Note that any sensor can be a destination; for example, a sensor may collect information from its n-hop neighbors (where n is significantly smaller than the network diameter) to maintain a local model of the environment.

- The separation theorem only applies in the limit of a large number of measurements. Since sensor networks often need only very low data rates, efficiency suffers due to short block lengths unless massive latency is acceptable. Perhaps more importantly, separation cannot be optimum—even asymptotically—in networks where the condition of multiple access exists [1].
- In multi-hop networks, the orthodox approach uses regenerative processing, whereby each hop is protected via a combination of forward error-correction coding and ARQ (retransmission) in case of decoder failure. While this allows higher-level protocols to assume error-free hop links, retransmissions are extremely expensive in terms of energy, and may be frequent in low-power sensor nets with poor-quality links.
- Sensor nodes may be naturally heterogeneous, with different computational capabilities and energy supplies. For example, a network may consist of numerous low-cost, low-capability sensors that simply sample and forward data to a few high-capability destinations with renewable energy sources, enabling inferential algorithms that have highly asymmetric computational requirements.

Herein we consider a signal-processing approach that exploits the redundancy inherent in both the source information from multiple sensors and the individual channel-coded sensor data. This redundancy is captured as prior information in terms of a probabalistic *global model* for the spatio-temporal processes measured at the source sensors, modified by noisy transducers and unreliable sensor nodes. Our primary goal is to determine if channel coding and a decoder equipped to jointly exploit channel codes and the global model can achieve higher energy efficiency, as expressed by lower error rates, than separate source and channel coding in a sensor network.

Each measurement is channel-coded individually at the source sensor. The collection of coded measurements is a *source-channel product code*. Consider the case of systematic block coding of the measurements. Then if the symbol-level representations of the coded measurements are stacked vertically, the channel codes protect the rows using deterministic (e.g, parity) constraints, while the redundancy of the data protects symbols both within, and more importantly, between measurements with probabilistic equality constraints encoded in the global model. In essence, each bit of a received measurement is both an information bit and a constraint bit, in that it provides information about its own datum as well as redundancy for other data it is correlated with.

The source sensors do not need knowledge of the global model. At the destination, it can be constructed from other data (e.g., a previously-deployed network in a similar environment) or developed as the data arrives. The destination must estimate the set of original variables (or signals) from the received collection of coded measurements. In our approach, channel and source decoding are not separated. The decoder uses reliability information from the source-destination channels in addition to the global model, blurring boundaries between error control and estimation. We describe a method to accomplish this called source-channel (SC) decoding, a Bayesian inference technique. To bracket the performance of any SC decoding algorithm, we consider two SC decoding algorithms: the MAP estimator, and a much simpler message-passing decoder wherein the error correction code and the global model are invoked iteratively to arrive at an estimate of the environment. We refer to this decoder as a product code message-passing (PCMP) decoder, since the two dimensions (row/channel and column/global model) are handled by separate decoders that exchange *a posteriori* symbol probabilities at each iteration. Product codes composed of component block codes which are iteratively decoded turbo-style using component MAP decoders were introduced in [2] as block turbo codes.

To provide a lower bound on SC decoding performance, the decoding algorithm is minimum-complexity in that only partial inter-measurement statistics are used. The PCMP decoder is also useful in that it can be modified to allow comparison of the performance of SC decoding with decoding of only one of the constituent codes.

Research activity in coding applicable to sensor networks has increased dramatically in the last decade. Much work (e.g., [3] and [4]) is based on existence results for the CEO problem [5], viz., transport of multiple measurements of a single source to a single destination. For correlated sources, it is shown in [6] that separate source coding, channel coding, and routing is asymptotically optimal for lossless reconstruction when the network consists of a set of independent point-to-point communication channels. More applicable to our work is the development of practical distributed source codes for correlated measurements (see the review [7]), motivated by existence results for lossless [8] and lossy [9] distributed source coding. The viewpoint of this paper was inspired by the exploration of lossy source-channel coding of two memoryless sources using turbo codes [10]. Use of the turbo principle in iterative source-channel decoding is investigated in [11] for variable-length source coding, and [12] for a single first-order correlated source, which is bit-interleaved and convolutionally-encoded for transmission over an AWGN channel.

In this paper, we design the source-channel decoder based on probabilistic constraints derived from the global model; this generalizes earlier code designs that employ binning strategies [13] based on deterministic Hamming-distance constraints. The dependence of our approach on the global model can be considered a disadvantage as it is specific to the application; however, it seems likely that exploitation of this knowledge will be very useful in practice given the importance of minimizing energy consumption and the limitation of block lengths. In many sensor network applications, decoding can be considered a special case of model-based inference of state variables and important parameters of the sensed processes. In these scenarios, the Hamming distance measure is clearly a poor match to the symbol-level correlations found in datasets. In contrast, the global model in this paper is a joint probability mass function over the symbols of the SC code that can be easily generated from any joint distribution over the sensed data.

II. MODEL

This paper is motivated by a range of practical wireless sensor networking scenarios that require the deployment of many sensors measuring multiple environmental processes with memory that are correlated over time and space. For example, in environmental sensing, light and temperature form a spatio-temporal vector process that is strongly correlated over space and time both within and between components. A collection of these measurements is acquired at the destination. The method of transport, whether one- or multi-hop, is arbitrary. In this model, any sort of single-hop error correction can be used along the way if desired; the destination only requires an estimate of the overall error probability of each source-destination channel.

Driven by financial constraints, radios in current sensor networks are usually simple, often using binary OOK or FSK modulation and hard-decision bit detection. Thus, while our setup is completely general in that it can use soft decisions, we model the overall channel from sensor i to the destination as binary symmetric with a probability of error ρ_i. Note that ρ_i can vary for each sensor; this is crucial since in real-world networks a channel tends to deteriorate with the number of its constituent single-hop links. The overall source-destination probability of error can be determined either prior to deployment or learned in the network, and in the latter case can be based on end-to-end measurements or on single-hop statistics and knowledge of the route.

In this paper, we consider in detail the use of low-complexity channel codes: a single parity-check code and a $(n, k, d_{\min}) = (7, 4, 3)$ Hamming code. While any code can in principle be used, we are interested in coding that does not require a large number of samples for efficiency and imposes little computational load on the reporting sensor.

Let x_i be the sample value of some physical process measured by a transducer of some sensor. Here i is a general index that subsumes sensor, transducer, and time indices. The measurement m_i of x_i is corrupted by transducer noise and then uniformly quantized into the information word u_i, which is transformed into a codeword v_i. The codeword is forwarded to the destination, which receives $y_i = v_i + z_i$ where z_i is a binary error vector with addition appropriately defined.

The entire collection of data samples is simply x, and similarly for m, u, v, and y. From Bayes' theorem, we have

$$p(x, m, u, v|y) \propto p(y|x, m, u, v)p(x, m, u, v), \quad (1)$$

which can be re-written as

$$p(x, m, u, v|y) \propto p(y|v)p(v|u)p(u|m)p(m|x)p(x) \quad (2)$$

where $p(x)$ is the prior information about the source process to be incorporated in the global model as $p(v, x)$ or as $p(u)$.

But the mapping of the per-measurement channel code from m to v is deterministic, so we can use

$$p(x, m, u, v|y) \propto p(y|v)p(v|x)p(x). \quad (3)$$

To simplify the presentation, we assume that channel errors are independent, so

$$p(y|v) = \prod_{i,j} p(y_{ij}|v_{ij}), \quad (4)$$

where y_{ij} denotes the j-th bit of y_i. Transducer noises are also assumed independent, yielding

$$p(v|x) = \prod_{i} p(v_i|x_i). \quad (5)$$

At the destination, the channel log-likelihood ratios (LLRs) $\lambda_{ij}^{\mathrm{ch}} = \log(p(y_{ij}|v_{ij} = 1)/p(y_{ij}|v_{ij} = 0)$ are found for each bit j of the received noisy coded sensor measurement y_i. For the binary symmetric channel (BSC) with crossover probability ρ considered in this paper, $p(y_{ij} = 1|v_{ij} = 1) = 1 - \rho$, $p(y_{ij} = 1|v_{ij} = 0) = \rho$, and similarly for $y_{ij} = 0$.

This paper emphasizes the effects of channel errors and hence focuses on inference of the collection of quantized measurements u. We thus re-define u to incorporate the properties of the source statistics and measurement noises in a symbol- or bit-level probabilistic model. The MAP estimate of the transmitted SC product codeword is then

$$\hat{u} = \arg\max_{u} p(u|y) = \arg\max_{u} p(y|u)p(u). \quad (6)$$

The global model is expressed as $p(v|x)p(x)$ in equation (3). In practice, the global model contains the conditional probabilities $p(v_{ij}|v)$ that include the correlation between bit v_{ij} and all other bits in v.

III. Decoder Structure

An upper bound on performance is obtained using a MAP decoder for the SC product code that searches over the set of all valid product codewords for the MAP estimate $\arg\max_v p(y|v)p(v)$, where $p(v)$ depends on the source statistics and is not uniform. This decoder may be practical in many scenarios. In sensor networks, even the aggregate data rate received at the destination may be very small; bandwidths of environmental signals are often less than 1 Hz. Often the destination has substantial computational and energetic capabilities far exceeding those of the sensor nodes, or the user may not require real-time receipt of the data.

The PCMP decoder is much simpler than the SC MAP decoder, as it employs two separate decoders, one for each dimension of the SC product code. The first is a channel decoder that implements the deterministic parity constraints of the per-measurement channel code. The second decoder is constructed from the probabilistic symbol-level constraints derived from the global model. The PCMP decoder is a turbo decoder; at each iteration, the decoders are invoked sequentially, generating *a posteriori* probabilities to be used as *a priori* information for the other decoder. Due to its modular structure, the PCMP decoder may be useful in some applications, since a large number of SC product code decoders can be quickly snapped together as needed from a collection of channel decoders and global model-derived inference engines. The global model contains the conditional probabilistic structure of the source model. This model may be made as precise as desired, assuming infinite storage resources.

There are two decoding elements within the PCMP decoder. The first element (top, Figure 1) consists of N channel decoders, one corresponding to each sensor. The channel LLRs $\lambda_{ij}^{\mathrm{ch}}$ are input to channel decoder i, along with any *a priori* information on v_i. The channel decoders are maximum *a posteriori* (MAP) decoders for the row codes, or individual channel codes, of the SC product code. They generate *a posteriori* output LLRs $\lambda_{ij}^{\mathrm{APP}}$, that are used as *a priori* input to the second element of the PCMP decoder, the global model.

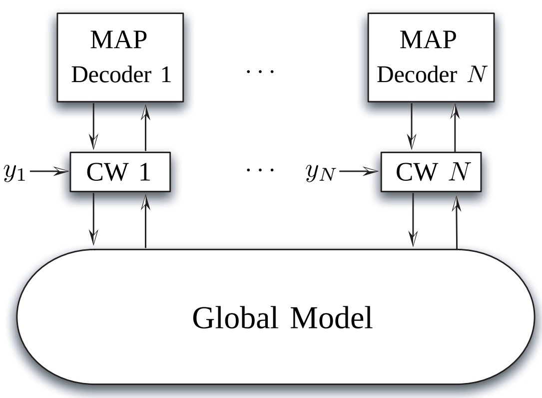

Fig. 1. Block Diagram of PCMP decoder, with component channel decoders and global source model.

The function of the global decoder (bottom, Figure 1) is to combine the *a priori* information on the codewords v_i from each channel decoder with the conditional probability information of the global model.

To provide a lower bound on performance and complexity, the PCMP decoder implementation of the global model uses bitwise information only, that is, the conditional probability of bit j of measurement i given bit j of all other measurements.

This bitwise simplification of the global model correlation is suboptimal; however, it significantly reduces the size of the conditional probability matrix. For N sensors and a codeword length of n bits, the SC product code v has Nn elements. The size of the required conditional probabilities is reduced from

$Nn * 2^{Nn}$ for full correlation to $Nn * 2^N$ when assuming bitwise correlation; this is a factor of 2^n reduction in size. For a (7,4,3) Hamming code, the reduction is a factor of 128. Since in our simulation experiments the global source model is symmetric, i.e., invariant with respect to the ordering of the measurements, these probabilities are symmetric as well. This further reduces the size of the required probabilities to $n * 2^N$.

The global model stores the bitwise conditional probabilities $p_{/i}(v_{ij}|v_{1...N,j})$ where $p_{/i}(v_{1...N,j})$ indicates that $v_{1...N}$ does not contain the ith measurement. From the channel decoders, it receives *a priori* $p(v_{ij})^{(l)}$ during iteration l.

Using the chain rule of conditional probability, and assuming bitwise correlation, the joint probability $p_{/i}(v_{1...N,j})$ is $p_{/i}(v_{1...N,j}) = p_{/i,1}(v_{1,j}|v_{2...N,j}) \dots p(v_{N-1,j}|v_{N,j}) \cdot p(v_{N,j})$. Assuming independence between bit j of all measurements, the joint probability $p_{/i}(v_{1...N,j})$ is the product of the individual probabilities,

$$p_{/i}(v_{1...N,j}) = \prod_{\substack{m=1 \\ m \neq i}}^{N} p(v_{m,j}). \tag{7}$$

The measurements are not independent; however, because the iterative decoder separates source and channel decoding, the *a priori* information from one channel decoder is not conditioned on information from the other decoders. To utilize the channel decoder LLRs as *a priori* information, we approximate the joint probability $p_{/i}(v_{1...N,j})$ according to equation 7.

From Bayes' theorem, the joint probability $p(v_{1...N,j})$ including v_{ij} is then approximated as

$$\begin{aligned} p(v_{1...N,j}) &= p_{/i}(v_{ij}|v_{1...N,j}) \cdot p_{/i}(v_{1...N,j}); \\ p(v_{1...N,j}) &\approx p_{/i}(v_{ij}|v_{1...N,j}) \cdot \prod_{\substack{m=1 \\ m \neq i}}^{N} p(v_{m,j}). \end{aligned} \tag{8}$$

The individual probabilities $p(v_{ij})^{(l+1)}$ are found by marginalization as

$$p(v_{ij})^{(l+1)} = \sum_{\substack{v_{1...N,j} \in V^{N-1} \\ v \neq v_{ij}}} p(v_{1...N,j}); \tag{9}$$

$$p(v_{ij})^{(l+1)} \approx \sum_{\substack{v_{1...N,j} \in V^{N-1} \\ v \neq v_{ij}}} p_{/i}(v_{ij}|v_{1...N,j}) \cdot \prod_{\substack{m=1 \\ m \neq i}}^{N} p(v_{m,j}).$$

These updated probabilities $p(v_{ij})^{(l+1)}$ or LLRs are sent to the channel decoders for use as *a priori* LLRs $\lambda_{ij}^{v,(l+1)}$ in iteration $l+1$.

One iteration consists of a complete exchange of information between the individual channel decoders and the global model. A decision on the information bits u_{ij} can be made at any time based on the MAP decision rule

$$\begin{aligned} \hat{u}_{ij} &= \arg\max_{u_{ij}} p(u_{ij}|y_i); \\ \hat{u}_{ij} &= \text{sign}(\lambda_{ij}^{\text{APP}}), \end{aligned} \tag{10}$$

for binary u_{ij} with APP LLRs $\lambda_{ij}^{\text{APP}}$.

IV. Simulation Results

In our simulation experiments, we use a symmetric multivariate Gaussian global source model, where the *a priori* information uncertainty is uniform over all measurements, i.e., the covariance matrix for the global model has diagonal elements $\sigma_{x_i}^2 = \Sigma_{ii} = \sigma_x^2$ for all i. The correlation r between measurements is determined from the information covariance, or off-diagonal elements $\sigma_{x_{ij}}^2 = \Sigma_{ij}$; the correlation is found as $r_{ij} = \Sigma_{ij}/\Sigma_{ii}$ for all $i \neq j$. The information mean is known *a priori*. Note that these measurements could result from a spatial snapshot across all sensors, a time series from one sensor, measurements of different parameters at the same sensor, or any combination thereof. The network of N sensors described in the simulations could represent N measurements taken at a single sensor.

For simplicity, we assume that transducer noise is Gaussian, implying that given x_i, m_i is normally distributed with mean x_i and variance σ_x^2, so that, due to the deterministic mapping from m to v,

$$p(v_i|x_i) = p(v_i = V_i|x_i) = \int_{R_{U_i}} f_m(s) ds \tag{11}$$

where $f_m(\cdot) \sim N(x_i, \sigma_x^2)$, V_i is the codeword corresponding to the quantized data U_i, and R_{U_i} is the range of the analog-to-digital converter input that is uniformly quantized into the value U_i. The tranducer noise variance is assumed zero.

Results were obtained for the SC product code using the MAP estimate, which views the SC product code as a single codeword, and the PCMP iterative decoder. Our measure of performance is the bit error rate (BER) of the estimated information bits $\hat{u}$ for each measurement. The channel is a binary symmetric channel (BSC) with crossover probability ρ ranging from 0.4 to 0.005, representative of the typically lossy channels encountered within the data-gathering sensor nets. For simplicity, ρ is identical for all transmitted measurements.

Two different low-complexity error-control codes are considered separately for use in the SC product code: the (4,3,1) single parity-check code (SPC) and the (7,4,3) Hamming code. Short block-length codes are used to demonstrate proof of concept. Performance would be expected to improve with longer channel codes. Both channel decoders use the BCJR algorithm operating on the parity-check trellis of the code to allow *a priori* information from the global model to be incorporated into its calculation of output *a posteriori* probabilities (APP).

Figure 2 displays the BER vs. BSC error probability ρ for the iterative PCMP decoder with a (4,3,1) SPC channel code. The network includes 6 sensors, with measurement variance $\sigma_x^2 = 0.2$ and measurement correlation $r_{ij} = 0.9$. Each measurement m_i is uniformly quantized to 3 information bits u_i and encoded by the (4,3,1) SPC code to 4 coded bits v_i.

The BER from the SPC portion of the PCMP decoder after 5 decoding iterations is compared to the BER from MAP decoding of the SC product code, using product codeword probabilities. The MAP estimate uses *a priori* $p(v)$, found from v generated from x according to the source statistics.

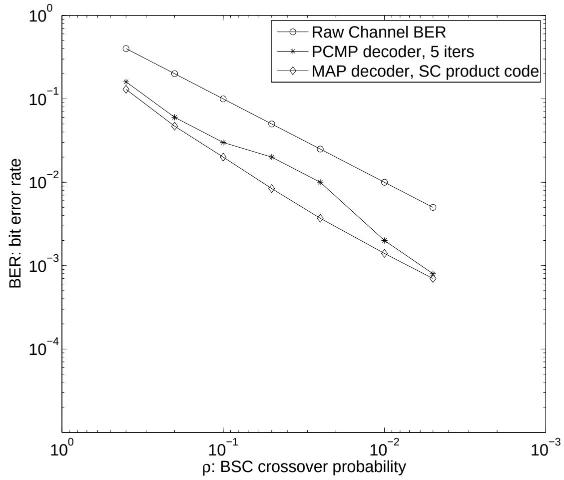

Fig. 2. Comparison of PCMP decoder, 5 iterations, using SPC channel code to MAP decoding of the SC product code; 6 sensors; source parameters $\sigma_x^2 = 0.2$ and $r_{ij} = 0.9$.

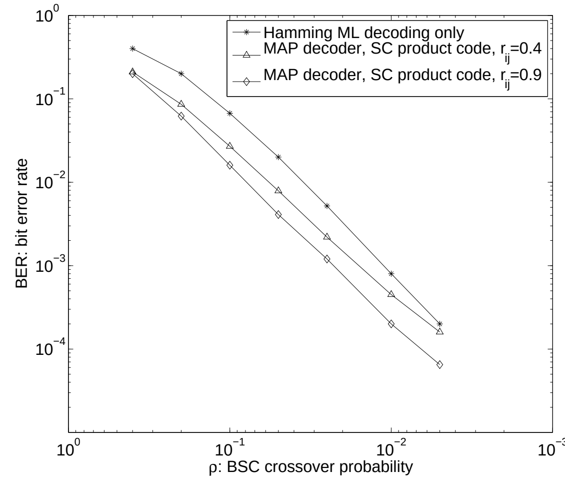

Fig. 3. Comparison of MAP decoding of SC Hamming product code using source statistics and ML decoding of individual (7,4,3) Hamming codes only; 3 sensors; source parameters $\sigma_x^2 = 1$, $r_{ij} = 0.4$ and 0.9.

The SPC alone is unable to correct any errors. However, in combination with the global model, the BER is reduced significantly below the channel probability of error ρ, and is close to that achieved by MAP decoding of the SC product codeword. Hence even though the per-measurement channel code is useless by itself for error correction, the SC product code exploits it effectively. This performance is as expected for an iterative source-channel decoder; the novelty is integration of the global model into the PCMP decoder as source decoder.

Use of a stronger error-correcting channel code should improve the performance of the PCMP iterative decoder, as well as that of the MAP decoder for the product code.

To underscore the performance improvement possible when incorporating probabilistic source knowledge in the decoder, the BER for MAP decoding over the entire $N \times 7$ SC product code using the product codeword probabilities $p(v)$ as *a priori* information is compared to ML decoding for each (7,4,3) Hamming code on the BSC in Figure 3. The network is composed of 3 sensors ($N = 3$).

The product codeword probabilities $p(v)$ depend on the global source probabilities. Highly correlated measurements produce a very narrow probability mass function centered about a few SC product codewords v; in fact, we have a near-repetition code, since nearly the same (7,4,3) Hamming codewords are repeated N times in an N-sensor network. (Note that the CEO problem, where the N sensors observe the same source with no measurement noise, is a special case.) Most product codewords will then have negligible probability mass. MAP decoding of the product code itself is thus computationally very feasible when the source information is highly correlated across measurements.

For poor-quality channels, we see a significant gap between the ML decoding performance of the Hamming code alone, and MAP decoding of the product code. The sub-optimum PCMP decoder also provides good performance improvement over ML decoding of the Hamming code alone. Thus SC product coding may be very useful in wireless sensor networks, where source-destination channels can be very poor. With increasing channel reliability, ML decoding of the (7,4,3) Hamming code alone approaches that of MAP decoding of the product code, for lower correlation between measurements. In good channels, use of the Hamming code alone is sufficient to achieve most of the performance gained by incorporating knowledge of the channel statistics into the decoding process, without the extra complexity. If the channel is known to be good, the iterative PCMP decoder need not be invoked.

Next we consider the iterative PCMP decoder using a (7,4,3) Hamming channel code. The network consists of three sensors, and the source statistics are as follows: the measurement variance $\sigma_x^2 = 1$ and correlation $r_{ij} = 0.4$. Each measurement m_i is uniformly quantized to four information bits u_i, then encoded by the Hamming code to seven coded bits v_i.

Figure 4 shows the BER versus ρ for the global decoder alone ($\circ$), for ML decoding of the (7,4,3) Hamming code only ($\times$), and for the Hamming decoder portion of the PCMP decoder after 5 iterations ($*$). MAP decoding of the product code ($\diamond$) using *a priori* knowledge of the product codeword probabilities is included as a bound on performance. Replacing the (4,3,1) SPC channel code with a (7,4,3) Hamming code does enhance the PCMP decoder performance.

The individual BER curves for each component code in Figure 4 shows that the two codes complement each other in their region of best performance; the (7,4,3) Hamming code performs better than the global source decoder in better-quality channels, while the global source decoder provides better performance in poor channels. Combining the two codes into an iterative PCMP decoder provides a performance improvement over either decoder alone in poor-quality channels. As channel quality improves, the performance of the (7,4,3) Hamming code alone converges to that of the PCMP decoder.

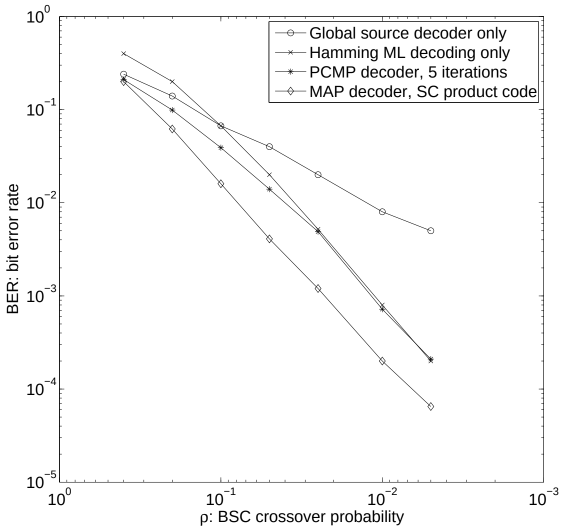

Fig. 4. Comparison of PCMP decoder, 5 iterations, with global source model only, ML decoding of (7,4,3) Hamming code only and MAP decoding of the SC product code; 3 sensors; source parameters $\sigma_x^2 = 1$, $r_{ij} = 0.9$.

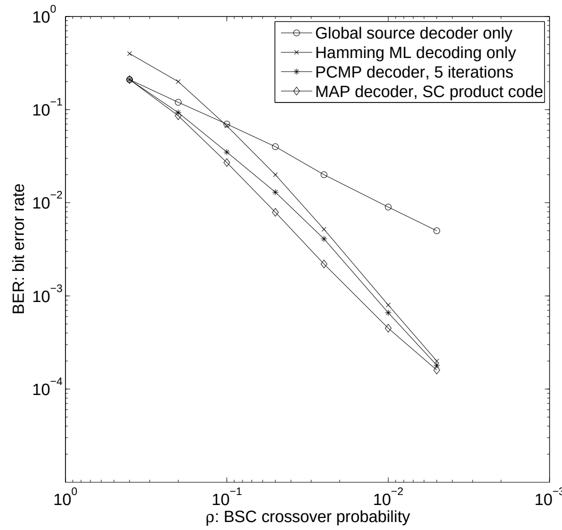

Fig. 5. Comparison of PCMP decoder, 5 iterations, with ML decoding of (7,4,3) Hamming code alone and MAP decoding of the SC product code; 3 sensors; source parameters $\sigma_x^2 = 1$, $r_{ij} = 0.4$.

The gap between the iterative PCMP decoding and the product code MAP decoding BER curves shows sub-optimality in the PCMP decoder. This sub-optimality is no surprise, as the iterative decoder breaks the decoding process into two component decoders instead of the single combined MAP decoder. This performance gap should diminish with use of a longer block-length code. Additionally, the global model considers only bitwise correlation, when there will be some correlation between adjacent bits from different sensors.

Figure 5 shows the BER versus ρ for the same PCMP decoder as Figure 4, again for three sensors, with $r_{ij} = 0.4$.

The PCMP decoder provides performance improvement compared to ML decoding of the (7,4,3) Hamming code alone, without any knowledge of the source probabilities, for poor-quality channels. As the channel quality improves, ML decoding approaches that of MAP decoding with the source probabilities, allowing little room for performance enhancement; in this case, ML decoding of the Hamming code alone converges to the PCMP decoder performance.

V. Conclusions

To minimize overall system lifecycle cost (including provision of energy), wireless sensor network designers employ sensor nodes with low-quality tranducers, minimal computational power and simple radios. In applications where the inference of evolving spatio-temporal processes is required, the probabilistic structure of the sensed data itself can be used as a network-wide global code that, when combined with simple error-control codes, can be exploited at the decoder to improve network energy efficiency, inferential performance, or both. Work is on-going to further develop these ideas in the context of inference of the state and important parameters of higher-level modelling of embedding environments.

Acknowledgments

The work of Dr. Howard was supported by an NAU Intramural Grant. The work of Dr. Flikkema was supported in part by NSF grant CNS-0540414 and the Center for Wireless Communications, University of Oulu.

References

[1] M. Gastpar and M. Vetterli, “Source-channel communication in sensor networks,” in *Proc. IPSN ’03*, 2003, pp. 162–177.
[2] R. M. Pyndiah, “Near-optimum decoding of product codes: block turbo codes,” in *IEEE Trans. Communications*, vol. 46, no. 8, Aug. 1998, pp. 1003–1010.
[3] P. Ishwar *et al.*, “On rate-constrained distributed estimation in unreliable sensor networks,” *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 4, pp. 765–775, Apr. 2005.
[4] K. Liu, H. E. Gamal, and A. Sayeed, “On optimal parametric field estimation in sensor networks,” in *Proc. IEEE Statistical Signal Processing Workshop*, 2005.
[5] T. Berger, Z. Zhang, and H.Viswanathan, “The CEO problem,” *IEEE Trans. Inform. Theory*, vol. 42, pp. 887–902, May 1996.
[6] J. Barros and S. Servetto, “Network information flow with correlated sources,” *IEEE Trans. Inform. Theory*, vol. 52, no. 1, pp. 155–170, Jan. 2006.
[7] Z. Xiong, A. Liveris, and S. Cheng, “Distributed source coding for sensor networks,” *IEEE Signal Processing Mag.*, pp. 80–94, Sep. 2004.
[8] D. Slepian and J. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inform. Theory*, vol. 19, pp. 471–480, Jul. 1973.
[9] A. Wyner and J. Ziv, “The rate distortion function for source coding with side information at the decoder,” *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
[10] J. Garcia-Frias and Y. Zhao, “Compression of binary memoryless sources using punctured turbo-codes,” *IEEE Communication Letters*, vol. 6, pp. 394–396, Sep. 2002.
[11] J. Hagenauer and R. Bauer, “The turbo principle in joint source-channel decoding of variable length codes,” in *Proc. IEEE Inform. Theory Workshop*, 2001, pp. 33–35.
[12] J. Hagenauer and N. Gortz, “The turbo principle in joint source-channel coding,” in *Proc. IEEE Inform. Theory Workshop*, 2003, pp. 275–278.
[13] S. Pradhan and K. Ramchandran, “Distributed source coding using syndromes (DISCUS): Design and construction,” *IEEE Trans. Inform. Theory*, vol. 49, no. 3, pp. 626–643, Mar. 2003.