

Optimal matching in wireless sensor networks

A. Roumy[†], D. Gesbert^{*}

[†] INRIA-IRISA, Rennes, France.

^{*} Institute Eurecom, Sophia Antipolis, France.

September 2007

Abstract

We design a wireless sensor network (WSN) in terms of rate and power allocation in order to send without loss the data gathered by the nodes to a common sink. Correlation between the data and channel impairments dictate the constraints of the optimization problem. We further assume that the WSN uses off-the-shelf compression and channel coding algorithms. More precisely source and channel coding are separated and Distributed Source Coding (DSC) is performed by pairs of nodes. This raises the problem of optimally matching the nodes. We show that under all these constraints the optimal design (including rate/power allocation and matching) has polynomial complexity (in the number of nodes in the network). A closed form solution is given for the rate/power allocation, and the matching solution is readily interpreted. For noiseless channels, the optimization matches close nodes whereas, for noisy channels, there is a tradeoff between matching close nodes and matching nodes with different distances to the sink. This fact is illustrated by simulations based on empirical measures. We also show that the matching technique provides substantial gains in either storage capacity or power consumption for the WSN wrt the case where the correlation between the nodes is not used.

1 Introduction

We consider a wireless sensor network (WSN) where spatially distributed sensors (or sensor nodes) gather data and send them to a common center (or sink) in order to monitor some physical or environmental phenomenon [1, 10]. A design issue for such a WSN is to maximize the network lifetime while dealing with low-cost sensors exhibiting limited

¹This work is supported by the Network of Excellence in Wireless Communications (NEWCOM), E. C. Contract no. 507325.

²Part of this work was presented at the IEEE International Symposium on Information Theory (ISIT) 2007 [13].

capabilities in terms of processing (computation capabilities, memory) and communication (power).

A naive approach consists in transmitting all the data measured by the sensors directly to the sink. This approach suffers two sub-optimality: (1) First due to spatial correlation between the measured data, the sufficient amount of data to transmit from the nodes to the sink can be reduced (from the sum of individual entropies to the joint entropy). Therefore taking into account the correlation between the nodes, communication power and spectral resource can be saved. (2) Second the Distributed Source Coding (DSC) (aka. Slepian Wolf coding) theorem [14] states, that this reduced amount of data can be sent without explicit cooperation between the nodes. Therefore using DSC techniques can save not only some resource (no communication between the nodes) but also some processing (data of other nodes are not processed at a node). More precisely, in DSC, the only knowledge required at each node is the rate at which this node needs to compress its data. Note that all the processing complexity is transferred to the sink, since to achieve optimal compression without encoding cooperation, joint decoding has to be performed.

Therefore, in the context of WSN, optimal strategies have been proposed in the literature based on DSC coding [5], where an optimal DSC coding for all the nodes is assumed. However most existing DSC schemes concern two correlated sources. First attempt to design codes for multiple (binary) sources has been proposed in [16] but it suffers some loss wrt to the optimal compression rate. Therefore it is also of interest to consider DSC coding for pairs of nodes. This strategy is referred to *pairwise DSC* in the following. Note that pairwise DSC is close to the idea of clustering nodes as in in-network aggregation [9]. There is however some difference between the two approaches: in in-network aggregation, nodes need to communicate their data to their neighbors, and then a decision whether to compress or concatenate the data is taken, based on the correlation between the data. In contrast, pairwise DSC avoids transmission between nodes, correlation measurement and strategy optimization at the nodes. The only information required at each node, is the power and rate at which to compress the data but not the global strategy. Another reason for considering pairwise DSC is its flexibility. A system designed with DSC for two sources is more flexible than a single multiple source code. This results from the existence of DSC codes (for two sources) that can operate at any rate in the Slepian-Wolf region and that can adapt to any change in correlation between the sources [17]. Therefore, this unique code can be embedded in each sensor, and the matching (performed at the sink) will take care of the changes in topology.

Therefore, we focus in this paper on pairwise DSC coding. This raises a new question: how to optimally match the paired nodes. We address this problem in two different communication scenarios. (i) Perfect node-sink channels. In that context, the goal is to maximize the storage capacity. (ii) Orthogonal noisy channels. In that case, source channel separation holds [2] and we optimize the compression rates and the node matching in order to minimize the total used power.

In this paper, we assume that the sink has full knowledge of the individual, and

pairwise joint entropies but also of the pairwise channel capacity regions. This is the case, when these quantities only depend of the internode and node-to-sink distances and when the sink has full knowledge of the network topology. If not, then these quantities have to be estimated through training sequences sent by the nodes.

Main contributions of the paper. First we model the design of a pairwise DSC scheme in a WSN. Then we show that the optimization of the pairwise DSC strategy can be separated into a matching problem and a pairwise rate-power control problem (that admits a simple closed-form solution). Using this separation, we obtain an optimization procedure of polynomial (in the number of nodes in the network) complexity. Finally, we show that the overall optimization can be readily interpreted. For noiseless channels, the optimization matches close nodes whereas for noisy channels, there is a tradeoff between matching close nodes and matching nodes with different distances to the sink.

2 Sensor network model and problem statement

We consider a network with N nodes all communicating to a single sink. Let \mathcal{N} be the set of sensor indices: $\mathcal{N} = \{1, \dots, N\}$. The data to be sent from node $i \in \mathcal{N}$ to the sink are modeled as the realizations of a discrete random variable denoted X_i taking its value in the alphabet \mathcal{X} . R_i [resp., P_i] denotes the rate (in bits per source symbol) [resp., the power] at which node i sends data.

The data of all the nodes are compressed without loss with a pairwise DSC scheme. More precisely the data are encoded separately at each node but decoded (jointly) at the sink, by pairs of nodes. This approach is suboptimal in comparison to the joint decoding of the data of all nodes but is motivated by the availability of efficient DSC schemes for two sources (see references in [8]). This raises the question of optimally partitioning the nodes into pairs, which can be modeled as the selection of an optimal *2-partition* defined below.

Definition 1. *A 2-partition \mathcal{P} is a partition of \mathcal{N} s.t. the cardinality of each subset is 2, except for a set that contains only one element if N is odd. An element of 2-partition is called a pair (even for the left alone node).*

Let \mathcal{S} denote the set of all possible 2-partitions.

Property 1. *The total number of 2-partitions is:*

$$|\mathcal{S}| = \begin{cases} (N-1)!! & \text{if } N \text{ is even} \\ (N)!! & \text{if } N \text{ is odd} \end{cases}$$

Proof. If N is even, $|\mathcal{S}|$ is the number of ways to choose $N/2$ disjoint pairs of items from N items, which is $(N-1)!! = (N-1)(N-3) \dots 5.3.1$ [15]. If N is odd, first a node is chosen to build the set with 1 element (N choices), then for the $N-1$ resulting nodes $(N-2)!!$ different matchings can be build. \square

The problem we address in this paper is how to send the data (measured by the sensors) to the sink without loss. In a pairwise DSC scheme, the design parameters are: the 2-partition, the compression rates and for noisy channels the powers used to send the data. The cost function is application and communication dependent. In the following, we consider different communication scenarios (noiseless and noisy) and define the related cost functions.

3 Perfect node-sink channels

In this section, we assume that each node can communicate directly to the sink and that the channels between each node and the sink are perfect. In this context, we want to maximize the storage capacity of each node (sensors and sink) without losing any information. Since pairwise DSC is used, the rates of each pair of nodes are constrained to lie in the so-called Slepian and Wolf region [14]: two nodes i, j can separately code their source symbols without loss of information if their compression rates (in bits per source symbol) belong to the Slepian and Wolf region SW_{ij} defined as:

$$SW_{ij} \triangleq \left\{ (R_i, R_j) : \begin{array}{l} R_i \geq H(X_i|X_j) \\ R_j \geq H(X_j|X_i) \\ R_i + R_j \geq H(X_i, X_j) \end{array} \right\} \quad (1)$$

Fig. 1 represents the set of rate pairs (R_i, R_j) for which lossless compression is possible. The region in dark grey corresponds to separated source coding (separated encoding and decoding). The SW_{ij} region corresponds to DSC (i.e. separated encoding but joint decoding) and includes the dark and light grey region. DSC allows smaller rates (shown in light grey): this reduces the total amount of data to be sent or stored for the same amount of information captured by the WSN.

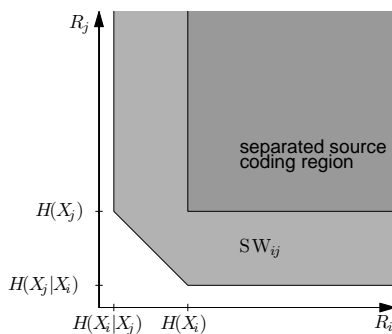


Figure 1: Slepian Wolf region

Hence, from the definition of the DSC rate region (1), the sensor network pairwise optimization for noiseless channels can be rewritten as Problem 1.

Problem 1. *The maximization of the storage capacity over all individual rates and over all 2-partitions under the constraint of lossless pairwise DSC coding reads:*

$$\begin{aligned} (\{R_i^*\}_{i=1}^N, \mathcal{P}^*) &= \arg \min_{\{R_i\}_i, \mathcal{P} \in \mathcal{S}} \sum_{i=1}^N R_i \\ \text{subject to} & \quad \forall (i, j) \in \mathcal{P}, \quad (R_i, R_j) \in \text{SW}_{ij} \end{aligned}$$

Since $\sum_{i=1}^N R_i = \sum_{(i,j) \in \mathcal{P}} R_i + R_j$, this problem can be reformulated as

$$\begin{aligned} (\{R_i^*\}_{i=1}^N, \mathcal{P}^*) &= \arg \min_{\{R_i\}_i, \mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} R_i + R_j \\ \text{subject to} & \quad \forall (i, j) \in \mathcal{P}, \quad (R_i, R_j) \in \text{SW}_{ij} \end{aligned}$$

At first sight, this problem looks exponentially hard due to the large number of possible 2-partitions (see Property 1). However, Proposition 1 and its Corollary 1 show that in fact it has polynomial complexity.

Proposition 1. *Separation of rate allocation and 2-partition selection. The storage capacity maximization (Problem 1) can be separated in a rate allocation over all possible distinct unordered pairs and a selection of the best partition. First, the rate allocation:*
 $\forall (i, j) \in \mathcal{N}^2$ s.t. $i < j$

$$R_{ij}^* = \arg \min_{R_{ij} \in \text{SW}_{ij}} R_{ij}(1) + R_{ij}(2) \quad (2)$$

where $R_{ij} = (R_i, R_j)$ denotes a rate pair and $R_{ij}(1) = R_i$ [resp., $R_{ij}(2) = R_j$] the first [resp., second] element of the pair. This rate allocation is independent of the 2-partition. Then, the 2-partition optimization:

$$\mathcal{P}^* = \arg \min_{\mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} R_{ij}^*(1) + R_{ij}^*(2) \quad (3)$$

Remark 1. *Motivation for the notation R_{ij} .* In order to justify the clumsy notation R_{ij} , we remark that the optimal value of R_i depends on i and on j . Therefore, we need to introduce a new notation for the pair of rates.

Proof. To solve problem 1, we can first minimize over the rates (for a given \mathcal{P}) and then over the 2-partition. This is always true as stated in [3, page 133], but note that, the rate minimization is performed for a given partition \mathcal{P} . In the following, we show that this rate allocation can be made independent of \mathcal{P} .

Part 1: for a given \mathcal{P} . First, consider the rate allocation problem for a given 2-partition \mathcal{P} :

$$\{R_i^*\}_{i=1}^N = \arg \min_{\{R_i\}_i \text{ s.t. } (R_i, R_j) \in \text{SW}_{ij}} \sum_{(i,j) \in \mathcal{P}} R_i + R_j$$

By definition of a 2-partition, a given rate R_i appears in only one constraint. Therefore the constraints are pairwise independent and the minimization can be carried out independently over each pair of \mathcal{P} [3, page 133]. This can be rewritten as:

$$\forall (i, j) \in \mathcal{P}, R_{ij}^* = \arg \min_{R_{ij} \in \text{SW}_{ij}} R_{ij}(1) + R_{ij}(2)$$

where $R_{ij} = (R_i, R_j)$ denotes the pair of rates and $R_{ij}(1) = R_i$ ($R_{ij}(2) = R_j$) an element of the pair. The minimum of the sum rate is therefore

$$\sum_{(i,j) \in \mathcal{P}} R_{ij}^*(1) + R_{ij}^*(2) \quad (4)$$

Part 2: for any \mathcal{P} . Now, for each \mathcal{P} , a rate allocation (as presented above) has to be performed, and the optimal 2-partition, solution of Problem 1, is the one that achieves minimum sum rate (4). In this procedure, rate allocation will have been performed for all possible unordered pairs. *Note that if the same pair appears in two distinct partitions, their rate allocation will be the same.* In other words, the rate allocation does not depend on the 2-partition chosen. Therefore we can first allocate the rates for all unordered pairs and then choose the best 2-partition (with minimum sum rate). This strategy is more efficient, since rate allocation for the same pair is made only once (and not each time the pair appears in a 2 partition). \square

Corollary 1. *The storage capacity maximization under lossless pairwise DSC coding is polynomial in the number of sensor nodes.*

Proof. From Proposition 1, the joint optimization separates into a rate allocation with complexity $O(N^2)$ and a partition optimization. The latter is a classical problem in combinatorial optimization, where it is known under the name of weighted matching for non-bipartite graph [11]. It can be solved in polynomial time [6]. More precisely, a straightforward implementation of Edmonds' algorithm runs in time bounded by $O(N^4)$, but more efficient implementations exist resulting in a bound of $O(N^3)$ [7]. A summary of the complexity results of Edmonds' algorithm is given in [4]. \square

The rate allocation can be further simplified. First, notice that the minimum of the sum $R_i + R_j$ in SW_{ij} is achieved for $R_i + R_j = H(X_i, X_j)$. It follows that partition selection can be reformulated as:

$$\mathcal{P}^* = \arg \min_{\mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} H(X_i, X_j) \quad (5)$$

Since (5) is independent of the rates, we can exchange the optimization order in Proposition 1. Therefore, the joint optimization can be achieved by first optimizing the 2-partition with respect to the joint entropies and then allocating the rates. This reduces the complexity of the rate allocation from $O(N^2)$ to $O(N)$. Further notice that the rate allocation is not unique. Any rate of the dominant face of the Slepian Wolf region solves Problem 1 (these rates are: (R_i^*, R_j^*) s.t. $R_i^* + R_j^* = H(X_i, X_j)$ and $R_i^* \geq H(X_i|X_j)$ and $R_j^* \geq H(X_j|X_i)$). It leads to the following algorithm:

Algorithm 1. *Design of a WSN (N nodes) with perfect node sink channels.*

- *Weight computation*
for $i = 1 : N$,

for $j = i + 1 : N$,
 $W(i, j) = H(X_i, X_j)$
 $W(j, i) = H(X_i, X_j)$
end
end

- *Weighted matching with weights $W(i, j)$ provides optimal 2-partition \mathcal{P}^* , that minimizes sum weight.*

- *Rate allocation for all pairs of the optimal 2-partition: \mathcal{P}^**
for $i = 1 : N/2$,
 $R_i = H(X_i)$
 $R_j = H(X_j|X_i)$, where j is the node matched to i in \mathcal{P}^*
end

4 Noisy node-sink channels

In this scenario, the channels between the nodes and the sink are noisy. More precisely, we assume independent additive white Gaussian noise (AWGN) channels. We also assume orthogonality in the channel access (no internode interference). This orthogonality can be achieved through protocols (Carrier Sense Multiple Access With Collision Avoidance (CSMA/CA)) or multiple access techniques (such as TDMA, FDMA or orthogonal CDMA). The capacity of the Gaussian channel (between node i and the sink) with transmit power P_i and channel gain γ_i is

$$C_i(P_i) \triangleq \log_2(1 + \gamma_i P_i)$$

where P_i represents the cost of sending C_i bits (per transmission) over the channel with gain γ_i , and where the noise power is normalized to one. Notice that the function $C_i(x) = \log_2(1 + \gamma_i x)$ depends on i upon γ_i . We further assume that the channel gains $\{\gamma_i\}_i$ are fixed quantities, known by the sink.

Due to power limitation at the sensors, the transmit power is constrained by a so called peak power constraint: $\forall i, P_i \leq P_{max}$. In this context, a natural cost function is the sum power that needs to be minimized. The constraints for this minimization are: the above mentioned peak power constraints and the asymptotically (in the size of the data length) small error probability.

We now detail the vanishing error probability constraint. Under the assumption of orthogonal channels, DSC and channel coding separation holds [2]. Therefore, the achievable (for vanishing error probability) rate region for distributed separated or joint source-channel coding¹ coincide. More precisely, for two sources, it is the set of rates (R_i, R_j)

¹*separated* refers to separation of source and channel coding, whereas *distributed* refers to separation of the processing between the sensor nodes.

lying in the intersection of the Slepian Wolf region SW_{ij} (1) and of the TDMA capacity region C_{ij} defined as:

$$C_{ij}(P_i, P_j) \triangleq \left\{ (R_i, R_j) : \begin{array}{l} R_i \leq C_i(P_i) \\ R_j \leq C_j(P_j) \end{array} \right\} \quad (6)$$

This achievable rate region (for distributed separated or joint source and channel coding) is the darker grey region on fig. 2. Hence, the sensor network pairwise optimization for

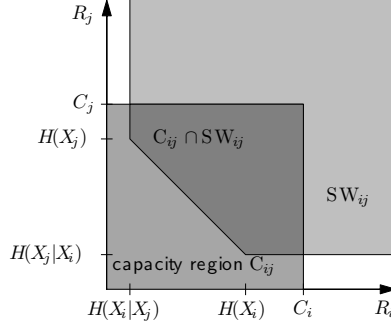


Figure 2: Slepian Wolf and capacity regions

orthogonal noisy channels can be stated as follows.

Problem 2. *The minimization of the transmit sum-power that achieves rates with vanishing error probability in a pairwise-distributed separated source and channel coding scheme² can be written as:*

$$\begin{aligned} (\{R_i^*\}_{i=1}^N, \{P_i^*\}_{i=1}^N, \mathcal{P}^*) &= \arg \min_{\{R_i\}_i, \{P_i\}_i, \mathcal{P} \in \mathcal{S}} \sum_{i=1}^N P_i \\ &= \arg \min_{\{R_i\}_i, \{P_i\}_i, \mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} P_i + P_j \\ \text{subject to} \quad &\forall (i, j) \in \mathcal{P}, \quad (R_i, R_j) \in SW_{ij} \cap C_{ij}(P_i, P_j) \\ &\forall i \in \mathcal{N}, \quad P_i \leq P_{max} \end{aligned}$$

Before we discuss the solution (see Section 4.2), let us first simplify the problem.

4.1 Optimization separation

Proposition 2. *Separation of rate-power allocation and 2-partition selection. The sum-power minimization (Problem 2) can be separated in:*

²Recall that for orthogonal channels, this scheme has same achievable rate region as the pairwise-distributed joint source and channel coding scheme

(i) a rate-power allocation over all possible distinct unordered pairs: $\forall (i, j) \in \mathcal{N}^2$ s.t. $i < j$

$$\begin{aligned} Q_{ij}^* &= \arg \min_{Q_{ij}} Q_{ij}(3) + Q_{ij}(4) & (7) \\ \text{subject to} & (Q_{ij}(1), Q_{ij}(2)) \in \text{SW}_{ij} \cap \text{C}_{ij}(Q_{ij}(3), Q_{ij}(4)) \\ & Q_{ij}(3) \leq P_{max} \\ & Q_{ij}(4) \leq P_{max} \end{aligned}$$

where $Q_{ij} = (R_i, R_j, P_i, P_j)$ denotes the four design parameters. The first two parameters of the quadruple are the rates $Q_{ij}(1) = R_i$, $Q_{ij}(2) = R_j$, whereas the last two represent the powers: $Q_{ij}(3) = P_i$, $Q_{ij}(4) = P_j$.

(ii) a matching (or 2-partition optimization):

$$\mathcal{P}^* = \arg \min_{\mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} Q_{ij}^*(3) + Q_{ij}^*(4) \quad (8)$$

Proof. Same proof as for Proposition 1. □

4.2 Sum of 2 powers minimization

In this section, we solve problem (7), a convex optimization problem of four variables. First, we show that the number of variables can be reduced to two (Lemma 1) and then to one (Lemma 2). In order to keep track of the meaning of the variables (which simplifies the proofs of the two lemmas), we shall use the notation R_i, R_j, P_i, P_j instead of Q_{ij} . This introduces no confusion, since the pair (i, j) is fixed in (7).

Lemma 1. *The minimum power is achieved on the boundary $P_i = \frac{2^{R_i}-1}{\gamma_i}$ [resp., $P_j = \frac{2^{R_j}-1}{\gamma_j}$].*

Proof. (7) is a convex optimization problem. The optimum occurs either at a stationary point or on the boundaries. Since there is no stationary point (linear function), the optimum occurs on a boundary s.t. P_i is minimum. It follows that $P_i^* = \frac{2^{R_i^*}-1}{\gamma_i}$. Similarly, we can show that $P_j^* = \frac{2^{R_j^*}-1}{\gamma_j}$. □

From Lemma 1, the rate-power optimization (ii) in Proposition 2 can be rewritten as:

$$\begin{aligned} Q_{ij}^* &= \arg \min_{Q_{ij}} \frac{2^{Q_{ij}^*(1)} - 1}{\gamma_i} + \frac{2^{Q_{ij}^*(2)} - 1}{\gamma_j} & (9) \\ \text{subject to} & (Q_{ij}(1), Q_{ij}(2)) \in \text{SW}_{ij} \cap \text{C}_{ij}(P_{max}, P_{max}) \end{aligned}$$

The following lemma allows to further reduce the number of variable.

Lemma 2. *The minimum power is achieved on the line $R_i + R_j = H(X_i, X_j)$.*

Proof. By contradiction. In the convex optimization problem (9), the convex objective function is strictly increasing and admits no stationary points. Therefore the optimum

occurs on a boundary.

We now show that the optimum can't occur on the line:

$$\begin{cases} R_i = H(X_i|X_j) \\ R_j \geq H(X_j|X_i) \\ R_j > H(X_i, X_j) - R_i = H(X_j) \geq H(X_j|X_i) \end{cases}$$

If $P_i^* > \frac{2^{R_i}-1}{\gamma_i}$ is the optimum, then there exists $\epsilon > 0$ s.t. $P_i^* = \frac{2^{R_i}-1}{\gamma_i} + \epsilon$. $P_i' = \frac{2^{R_i}-1}{\gamma_i}$ is in the region and $P_i' < P_i^*$ which contradicts the fact that P_i^* is minimum.

A similar argument holds for the region $R_j = H(X_j|X_i)$ and $R_i > H(X_i, X_j) - R_j = H(X_i)$.

Finally, the optimum occurs on the last boundary $R_i + R_j = H(X_i, X_j)$. \square

Finally, the rate-power allocation can be reformulated as a convex optimization problem of one variable:

$$\begin{aligned} Q_{ij}^*(1) &= \arg \min_{R_i} \frac{2^{R_i} - 1}{\gamma_i} + \frac{2^{H(X_i, X_j) - R_i} - 1}{\gamma_j} \\ \text{subject to} \quad &\text{lb} \leq R_i \leq \text{ub} \end{aligned} \quad (10)$$

where

$$\begin{aligned} \text{ub} &\triangleq \min \left(H(X_i), C_i(P_{max}) \right) \\ \text{lb} &\triangleq \max \left(H(X_i|X_j), H(X_i, X_j) - C_j(P_{max}) \right) \end{aligned}$$

From $Q_{ij}^*(1)$, all other variables can be deduced:

$$\begin{aligned} Q_{ij}^*(2) &= H(X_i, X_j) - Q_{ij}^*(1) \\ Q_{ij}^*(3) &= \frac{2^{Q_{ij}^*(1)} - 1}{\gamma_i} \\ Q_{ij}^*(4) &= \frac{2^{Q_{ij}^*(2)} - 1}{\gamma_j} \end{aligned}$$

Moreover, (10) admits a closed form explicit solution:

$$Q_{ij}^*(1) = \begin{cases} \text{lb} & \text{if } r < \text{lb} \\ r & \text{if } \text{lb} \leq r < \text{ub} \\ \text{ub} & \text{if } r \geq \text{ub} \end{cases} \quad (11)$$

where

$$r \triangleq \frac{1}{2} \left(H(X_i, X_j) + \log_2 \frac{\gamma_i}{\gamma_j} \right)$$

Therefore the algorithm that solves Problem 2 is:

Algorithm 2. Design of a WSN with noisy node sink channels and channel gains $\{\gamma_i\}_i$.

- *Weight computation*
for $i = 1 : N$,

for $j = i + 1 : N$,

$$r = \frac{1}{2} \left(H(X_i, X_j) + \log_2 \frac{\gamma_i}{\gamma_j} \right)$$

$$\text{ub} = \min(H(X_i), C_i(P_{max}))$$

$$\text{lb} = \max(H(X_i|X_j), H(X_i, X_j) - C_j(P_{max}))$$

$$\text{if}(r < \text{lb}) \quad W(i, j) = \frac{2^{\text{lb}} - 1}{\gamma_i} + \frac{2^{H(X_i, X_j) - \text{lb}} - 1}{\gamma_j}$$

$$\text{elseif}(r < \text{ub}) \quad W(i, j) = \frac{2^r - 1}{\gamma_i} + \frac{2^{H(X_i, X_j) - r} - 1}{\gamma_j}$$

$$\text{else}(r < \text{lb}) \quad W(i, j) = \frac{2^{\text{ub}} - 1}{\gamma_i} + \frac{2^{H(X_i, X_j) - \text{ub}} - 1}{\gamma_j}$$

$$W(j, i) = W(i, j)$$

end

end

- *Weighted matching with weights $W(i, j)$ provides optimal 2-partition \mathcal{P}^* , that minimizes sum weight. Minimal sum power is this sum weight.*

Solution interpretation. This results admits a nice interpretation. First, we rederive the rate-power optimization. More precisely, we seek for the optimal value of the sum power using a convexity argument. Since, the function $x \mapsto 2^x$ is convex, we have $\forall 0 \leq \alpha \leq 1$

$$2^{\alpha(R_i - \log_2 \gamma_i) + (1-\alpha)(R_j - \log_2 \gamma_j)} \leq \alpha 2^{R_i - \log_2 \gamma_i} + (1 - \alpha) 2^{R_j - \log_2 \gamma_j}$$

In particular, for $\alpha = \frac{1}{2}$, we have

$$\begin{aligned} & 2 \cdot 2^{\frac{1}{2}(R_i + R_j - \log_2 \gamma_i - \log_2 \gamma_j)} \\ \stackrel{\text{a}}{=} & 2 \cdot 2^{\frac{1}{2}(H(X_i, X_j) - \log_2 \gamma_i - \log_2 \gamma_j)} \\ \leq & 2^{R_i - \log_2 \gamma_i} + 2^{R_j - \log_2 \gamma_j} \end{aligned}$$

where (a) follows from the constraint $R_i + R_j = H(X_i, X_j)$. Therefore, the minimum of the cost function is achieved for

$$\begin{aligned} R_i^* &= \frac{1}{2} \left(H(X_i, X_j) + \log_2 \frac{\gamma_i}{\gamma_j} \right) \\ R_j^* &= \frac{1}{2} \left(H(X_i, X_j) + \log_2 \frac{\gamma_j}{\gamma_i} \right) \end{aligned} \tag{12}$$

which is equivalent to the second case in (11). But due to rate constraints (Slepian Wolf and capacity region) and peak power constraints, this optimal rate may not be achievable. In this case, bounds as lb and ub have to be introduced.

Therefore, due to the convexity of the exponential function, the optimal rate allocation (if feasible) corresponds to a balanced rate (up to the gain channel). More precisely, if

$\gamma_i = \gamma_j$, then $R_i = R_j = \frac{1}{2}H(X_i, X_j)$. Moreover, if these rates are feasible, the optimal powers are

$$\begin{aligned} P_i^* &= \sqrt{\frac{2^{H(X_i, X_j)}}{\gamma_i \gamma_j}} - \frac{1}{\gamma_i} \\ P_j^* &= \sqrt{\frac{2^{H(X_i, X_j)}}{\gamma_i \gamma_j}} - \frac{1}{\gamma_j} \end{aligned} \quad (13)$$

and the sum power is

$$2\sqrt{\frac{2^{H(X_i, X_j)}}{\gamma_i \gamma_j}} - \frac{1}{\gamma_i} - \frac{1}{\gamma_j} \quad (14)$$

Case without peak power constraint. Notice that the solution detailed above encompasses the case without peak power constraint by letting P_{max} tends to $+\infty$. Let us assume wlog that

$$\frac{2^{H(X_i|X_j)}}{\gamma_i} \geq \frac{2^{H(X_j|X_i)}}{\gamma_j}$$

the optimal rate reads

$$Q_{ij}^*(1) = H(X_i|X_j) + \left[\frac{1}{2} \left(H(X_i, X_j) + \log_2 \frac{\gamma_i}{\gamma_j} \right) - H(X_i|X_j) \right]^+$$

where $[x]^+$ denotes $\max(x, 0)$. The quadruple Q_{ij} is completely determined with (11).

4.3 Solution and complexity

We now consider the global optimization problem 2.

Corollary 2. *The sum power minimization under lossless pairwise-distributed separated source and channel coding is polynomial in the number of sensor nodes.*

Proof. From Proposition 2, the joint optimization separates into a power-rate allocation with complexity $O(N^2)$ (see Section 4.2) and a partition optimization with complexity $O(N^3)$ [4] (see Corollary 1). \square

Grouping N nodes? We note that the three dimensional matching problem is NP-complete [11, Theorem 15.7]. Therefore, grouping more than two nodes will lead to high complexity solution. This also explains why we restrict ourselves to the grouping of pairs of nodes.

Solution interpretation (continued). We continue the solution interpretation started in Section 4.2 and focus on the matching solution. For the sake of clarity, let us assume that the rates (12) are feasible (i.e. lie in the intersection of the capacity and SW region). The sum of two powers is given by (14). Therefore the 2-partition optimization matches pairs in order to minimize:

$$\mathcal{P}^* = \arg \min_{\mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} \sqrt{\frac{2^{H(X_i, X_j)}}{\gamma_i \gamma_j}} \quad (15)$$

In general, the joint entropy $H(X_i, X_j)$ decreases with the internode distance. Therefore, the minimization of an increasing function of joint entropies would rather match close pairs. On the other hand, the minimization of $\sum_{(i,j) \in \mathcal{P}} \frac{1}{\sqrt{\gamma_i \gamma_j}}$ tends to match pairs with different γ (see Corollary 4). Since the channel gain γ depends on the distance between a node and the sink, it follows that nodes that have different distances to the sink, will be matched.

The overall optimization (15) is therefore a tradeoff between matching close nodes and matching nodes with different distances to the sink. This optimization results in a pairing along radii from the sink. This fact is fully illustrated by our experiment later.

Theorem 1. *Let $\{\nu_i\}_i$ be a set of n ordered positive reals s.t. $\nu_1 \geq \nu_2 \geq \dots \geq \nu_n \geq 0$. Suppose n is even. Let \mathcal{P} be a 2-partition of n elements and \mathcal{S} the set of all possible 2-partitions. We have that*

$$\min_{\mathcal{P} \in \mathcal{S}} \sum_{(i,j) \in \mathcal{P}} \nu_i \nu_j = \sum_{i=1}^{n/2} \nu_i \nu_{n+1-i}$$

In other words, the optimal 2-partition matches most distinct reals.

Proof. By induction, it is sufficient to show that ν_1 and ν_n have to be matched in the partition that minimizes the sum weight. This argument is shown by contradiction.

Assume that there exists $i < n$ and $j > 1$ s.t. ν_i is matched with ν_1 and ν_j with ν_n in the perfect matching. From the ordering of the $\{\nu_i\}_i$, we have that $\nu_i - \nu_n \geq 0$ and $\nu_1 \geq \nu_j$. It follows that $(\nu_i - \nu_n)\nu_1 \geq (\nu_i - \nu_n)\nu_j$, which is equivalent to

$$\nu_1 \nu_i + \nu_j \nu_n \geq \nu_1 \nu_n + \nu_i \nu_j$$

This contradicts the fact that 1 is matched with i , and that j is matched with n in the optimal 2-partition. \square

Theorem 1 can be reformulated as:

Corollary 3. *Consider a complete, undirected and non-bipartite graph. Each node is given a positive weight, and the weight of an edge is the product of the weights of the nodes it is connected to. The perfect weighted matching problem (that minimizes the sum weight) admits a closed form expression.*

Corollary 4. *The minimization of $\sum_{(i,j) \in \mathcal{P}} \frac{1}{\sqrt{\gamma_i \gamma_j}}$ tends to match pairs with different γ .*

Proof. Apply Theorem 1 with $\nu_i = 1/\sqrt{\gamma_i}$. \square

4.4 Network Lifetime maximization

Up to now, we have addressed the total power minimization problem in a WSN. However, in some contexts (multihop scenario for instance), one can prefer to maximize the lifetime of the network, or equivalently minimize the maximum power (maximum taken over all the sensors of the network). Hence, network lifetime maximization for orthogonal noisy channels can be stated as follows.

Problem 3. *The minimization of the transmit max-power that achieves rates with vanishing error probability in a pairwise-distributed separated source and channel coding scheme can be written as:*

$$\begin{aligned}
\left(\{R_i^\Delta\}_{i=1}^N, \{P_i^\Delta\}_{i=1}^N, \mathcal{P}^\Delta \right) &= \arg \min_{\{R_i\}_i, \{P_i\}_i, \mathcal{P} \in \mathcal{S}} \max_{i \in \mathcal{N}} P_i \\
&= \arg \min_{\{R_i\}_i, \{P_i\}_i, \mathcal{P}} \max_{(i,j) \in \mathcal{P}} (\max(P_i, P_j)) \\
\text{subject to} \quad &\forall (i, j) \in \mathcal{P}, \quad (R_i, R_j) \in \text{SW}_{ij} \cap \text{C}_{ij}(P_i, P_j) \\
&\forall i \in \mathcal{N}, \quad P_i \leq P_{\max}
\end{aligned}$$

As for the sum-power minimization (see Problem 2), the lifetime maximization can be efficiently simplified.

Proposition 3. *Separation of rate-power allocation and 2-partition selection. The max-power minimization (Problem 2) can be separated in:*

(i) *a rate-power allocation over all possible distinct unordered pairs: $\forall (i, j) \in \mathcal{N}^2$ s.t. $i < j$*

$$\begin{aligned}
W_{ij}^\Delta &= \min_{Q_{ij}} \max(Q_{ij}(3), Q_{ij}(4)) \tag{16} \\
\text{subject to} \quad &(Q_{ij}(1), Q_{ij}(2)) \in \text{SW}_{ij} \cap \text{C}_{ij}(Q_{ij}(3), Q_{ij}(4)) \\
&Q_{ij}(3) \leq P_{\max} \\
&Q_{ij}(4) \leq P_{\max}
\end{aligned}$$

where $Q_{ij} = (R_i, R_j, P_i, P_j)$ denotes the four design parameters, and where W_{ij} is the weight to be used for weighted matching.

(ii) *a matching (or 2-partition optimization):*

$$\mathcal{P}^\Delta = \arg \min_{\mathcal{P} \in \mathcal{S}} \max_{(i,j) \in \mathcal{P}} W_{ij}^\Delta \tag{17}$$

Proof. Since $\sum_{i=1}^N P_i = \sum_{(i,j) \in \mathcal{P}} P_i + P_j$ and $\max_{i \in \mathcal{N}} P_i = \max_{(i,j) \in \mathcal{P}} (\max(P_i, P_j))$, sum-power and max-power minimization are similar problems. Therefore, the same proof as for Proposition 1 applies. \square

Polynomial complexity? To the best of our knowledge there exists no algorithm to solve the minimum max-weight matching problem, but we conjecture that it is also polynomial. Hence, we conjecture that the max-power minimization under lossless pairwise-distributed separated source and channel coding is polynomial in the number of sensor nodes.

We now analyze the solution of problem (16). As in section 4.2, we shall use the notation R_i, R_j, P_i, P_j instead of Q_{ij} , in order to keep track of the meaning of the variables.

Lemma 3. *The minimum max power (16) is achieved for $P_i^\Delta = P_j^\Delta$.*

Proof. Let E denotes the set of feasible (R_i, R_j, P_i, P_j) defined by the constraints in (16). The next trick consists in defining the two sets: $E_1 = E \cap \{(R_i, R_j, P_i, P_j) : P_i \leq P_j\}$ and $E_2 = E \cap \{(R_i, R_j, P_i, P_j) : P_i \geq P_j\}$. Since $E = E_1 \cup E_2$, we have that $\min_E = \min(\min_{E_1}, \min_{E_2})$. More precisely, we have

$$\begin{aligned} \min_{E_1} &= \min_{E \cap \{P_i \geq P_j\}} \max(P_i, P_j) \\ &= \min_{E \cap \{P_i \geq P_j\}} P_i \\ &= \min_{E \cap \{P_i = P_j\}} P_i \\ &= \min_{E \cap \{P_i = P_j\}} P_j \end{aligned}$$

The second equality follows from the fact that $\max(P_i, P_j) = P_i$, when $P_i \geq P_j$. The fourth equality is due to the fact that in the constraint $P_i = P_j$. The third equality can be shown by contradiction, as detailed below.

Proof for the third equality. By contradiction. Let $(R_i^\circ, R_j^\circ, P_i^\circ, P_j^\circ) = \arg \min_{E_1} \max(P_i, P_j)$. Assume that $P_i^\circ > P_j^\circ$. Now, $(R_i^\circ, R_j^\circ, P_i^\circ, P_j^\circ) \in E_1 \Rightarrow (R_i^\circ, R_j^\circ, P_i' = P_j^\circ, P_j^\circ) \in E_1$. Since $P_i' = P_j^\circ < P_i^\circ$, it contradicts the fact that P_i° is minimum.

Similarly, we have that $\min_{E_2} = \min_{E \cap \{P_i = P_j\}} P_j$. Since $\min_{E_1} = \min_{E_2}$, we get that

$$\min_E = \min \left(\min_{E_1}, \min_{E_2} \right) = \min_{E_1} = \min_{E \cap \{P_i = P_j\}} P_i$$

Therefore, the optimum is achieved for $P_i = P_j$. \square

Solution interpretation (continued). As intuition would suggest, the max-power minimization results in allocating equal powers (see Lemma 3). However for the sum-power minimization (see paragraph ‘‘Solution interpretation’’ in section 4.3), the optimal rate/power allocation will result in equal power (13) iff the matched nodes have equal channel gain $\gamma_i = \gamma_j$. Hence if nearest neighbors are matched (case of transmission of high data rates: $2^{H(X_i, X_j)} \gg \gamma_i \gamma_j$), max-power and sum-power minimizations will lead to similar results.

5 A sensor network example

In order to illustrate, the gain achieved by a pairwise sensor network optimization, we consider a WSN in a bounded square area where sensors are randomly placed. A sink is either placed at a corner or at the center. We consider AWGN channels between the nodes and the sink and assume that the channel gain is inversely proportional to the square distance between the nodes communicating together.

We consider the entropy model of [12] based on empirical measure of daily rainfall precipitation. All individual entropies are assumed equal. The joint entropy of two sources is a function of the individual entropy, of a coefficient c that captures the correlation, and

of the distances d_{ij} between the sources:

$$H(X_i, X_j) = H(X_i) + \left(1 - \frac{1}{1 + \frac{d_{ij}}{c}}\right) H(X_i) \quad (18)$$

The greater c , the more correlated the sources are.

5.1 Numerical examples

The sensors lie in a bounded area such that the coordinates of each sensor are in $[0, 1]^2$.

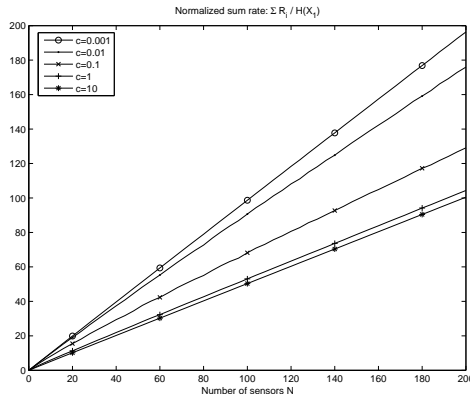


Figure 3: Perfect node-sink channel case. The normalized sum rate ($\sum_{i=1}^N R_i / H(X_1)$) as a function of the number of sensors. Result of the sum rate minimization, Problem 1, for different correlations c . Curves averaged over 10 realizations of the sensor positions.

Reduction of compression sum rate and sum power. In the case of perfect node-sink channel, Figure 3 shows the total sum rate needed to send all the data gathered by the sensors to the sink. Since in our example the sum rate scales multiplicatively with the individual entropy $H(X_i) = H(X_1), \forall i$ (18), we plot the normalized sum rate $\sum R_i / H(X_1)$. It is shown that the pairwise optimization helps to reduce the sum rate up to half of the naive case where all nodes communicate to the sink (in this case $\sum R_i = N \cdot H(X_1)$). Interestingly, the sum rate presents a smooth behavior, whereas it is averaged over only 10 realizations of the sensor placement. This may be explained by the fact that the weighted matching algorithm tends to match pairs with similar distances.

Figure 4 presents the total sum power consumed as a function of the number of sensors in the case of noisy channels. In this setup, we assume that the channel gain equals the inverse square distance between the node and the sink. Actually there is no need to add a scaling factor in the channel gain, because the rate allocation is insensitive to a scaling factor in the channel gain (see (12)) and the sum power will then scales multiplicatively with this constant (10). The reduction of sum power is even more important than the reduction of rate. This is due to the exponential behavior of the power wrt to the rate.

Matching result. The coefficient c captures the correlation. We now choose $c = 1$, such that if 2 sensors are distant by 1 (2 corners of the area), then the joint entropy is $\frac{3}{2}$ the

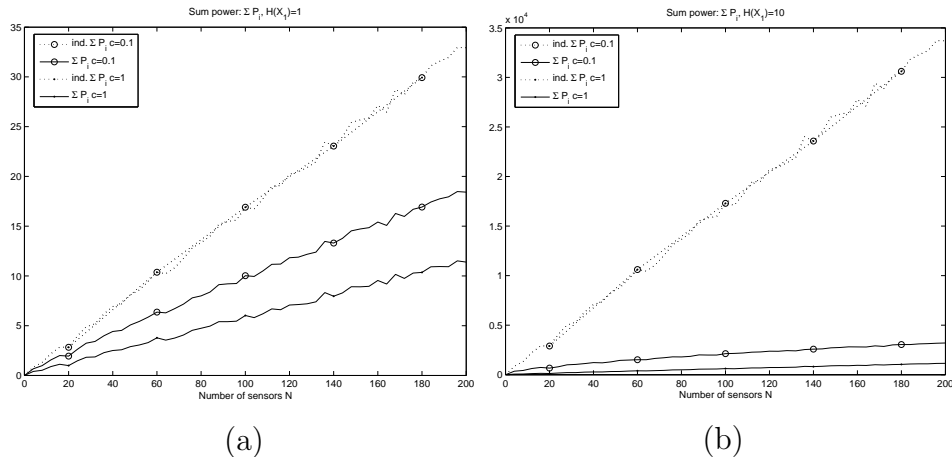


Figure 4: Noisy node-sink channel case. The total power consumed in the sensor network ($\sum_{i=1}^N P_i$) as a function of the number of sensors. Result of the sum power minimization, Problem 2, for different correlations c and different individual entropies $H(X_1)$. Curves averaged over 10 realizations of the sensor positions. (a) $H(X_1) = 1$ (b) $H(X_1) = 10$.

individual entropy. We therefore consider a highly correlated sensor network, that helps reducing the sum rate by about half (see fig. 3). Figure 5 shows the matching results for 20 sensors and a sink placed at the center of the area. The individual entropy is set to 1: $\forall i, H(X_i) = 1$. The numerical results (in terms of rate or power) are given in the title of each figure. As seen before, there is little variance of the sum rate and sum power wrt to the realization of the sensor placement. Therefore these values are significant.

For perfect channels, the figure that matters is the entropy between two sources or equivalently the distance. The optimization rather matches closest neighbors (see Fig. 5.a and 6.a), which is very intuitive since the joint entropy $H(X_i, X_j)$ decreases with the internode distance.

Once the power comes into the picture, the matching result changes and comparing Fig. 5 (a) and (b) shows that matched pairs are located on radius emanating from the sink. Fig. 6 and 7 highlight this observation. (On Figure 6 the sink is placed at the left bottom corner of the area, whereas it is at the center on Figure 7.) This can be further explained and understood with the help of the interpretation we gave in section 4.3. In these figures, we decrease the value of the entropy, s.t. the matching is more due to the influence of the γ and therefore distant nodes are matched together.

Effect of the peak power constraint. It is also interesting to note that adding a peak power constraint does not change the matching result (compare Fig. 5 (b) and (c)). However it can make the problem infeasible.

Maximum power. In the case of noisy channels, Figure 8 shows the distribution of the powers needed to send all the data to the sink. We compare the naive case where the nodes communicate directly to the sink (solid line) with the case, where the rate/power allocation and matching are done in order to minimize the sum-power. It is shown that the pairwise can significantly reduce the maximum power needed, especially when the

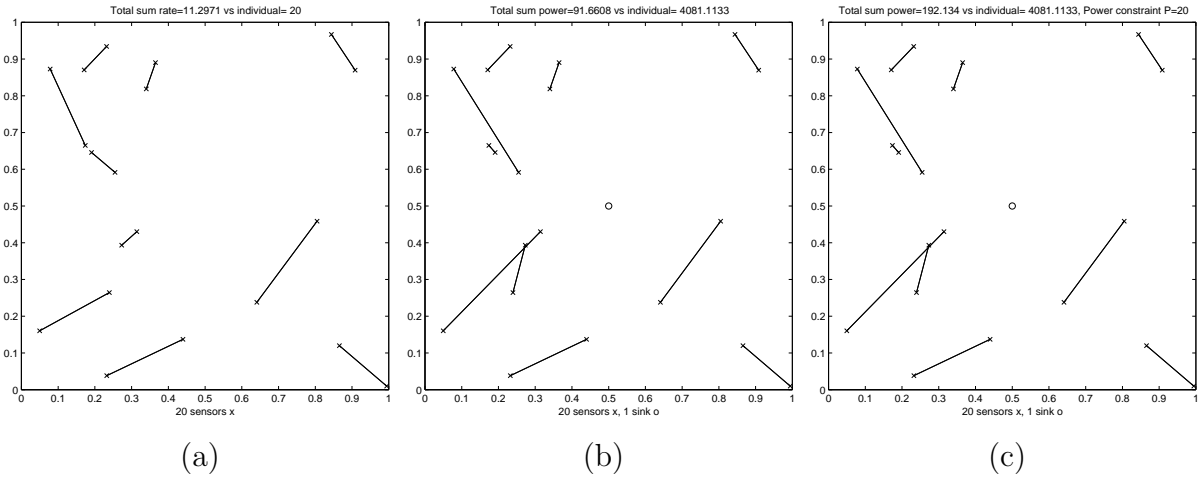


Figure 5: WSN with 20 sensors (x) and 1 sink (o) placed at the center. 2 matched nodes are linked by a line. Matching results for (a) perfect channel, (b) noisy channel without peak power constraint, (c) noisy channel with peak power constraint.

data are highly correlated ($c = 1$). When the individual entropy is high (see Figure 8(b) $H(X_1) = 10$), the sum-power minimization tends to match nearest neighbors (see Figure 7(b)). In that case, max-power minimization and sum-power minimization have similar rate/power allocation (as explained in section 4.4) and the reduction of the maximum power is even more significant.

6 conclusion

We proposed a way to design wireless sensor network while using existing and low complexity source and channel coding techniques. More precisely, only the correlation between pairs of nodes was taken into account. In comparison with the global DSC scheme (that processes the correlation between all the nodes), the complexity of our scheme is transferred to the sink. Indeed, an additional task is performed at the sink: it has to match the nodes into pairs. But, we showed that the complexity of the optimal design (including matching and rate/power allocation) is only polynomial in the number of nodes in the network. Finally, numerical results showed that the pairwise strategy can save about half the amount of data to be sent wrt the case where all the nodes transmit directly to the sink.

References

- [1] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. “Wireless sensor networks: a survey.” *Computer Networks*, 38(4):393422, April 2002.

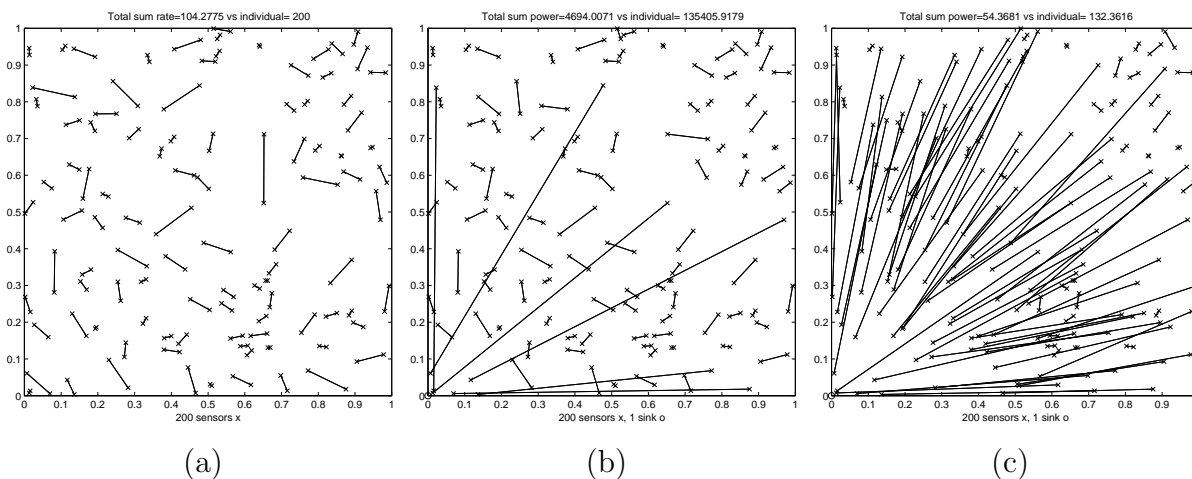


Figure 6: WSN with 200 sensors (x) and 1 sink (o) placed at a corner. 2 matched nodes are linked by a line. Matching results for (a) perfect channel. Then, noisy channel without peak power constraint, and individual entropy is 10 in (b), and 1 in (c).

- [2] J. Barros, S. D. Servetto, “Network Information Flow with Correlated Sources,” *IEEE Transactions on Information Theory*, Vol. 52(1), pp. 155-170, 2006.
- [3] S. Boyd, and L. Vandenberghe, *Convex optimization*, Cambridge University Press, 2004.
- [4] W. Cook and A. Rohe, “Computing Minimum-Weight Perfect Matchings,” *INFORMS Journal on Computing*, Vol. 11, pp. 138-148, 1999.
- [5] R. Cristescu, B. Beferull-Lozano, M. Vetterli, “Networked Slepian-Wolf: theory, algorithms, and scaling laws”, *IEEE Transactions on Information Theory*, 51(12):4057-4073, Dec. 2005.
- [6] J. Edmonds, “Paths, Trees and Flowers,” *Canadian Journal of Mathematics* 17, pp. 449-467, 1965.
- [7] H.N. Gabow, “Data Structures for Weighted Matching and Nearest Common Ancestors with Linking,” in *Proceedings of the First Annual ACM-SIAM Symposium on Discrete Algorithms*, Association for Computing Machinery, New York, 434443, 1990.
- [8] J. Garcia-Frias, Zixiang Xiong, “Distributed source and joint source-channel coding: from theory to practice,” *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '05)*, March 2005.
- [9] L. Krishnamachari, D. Estrin, S. Wicker, “The impact of data aggregation in wireless sensor networks,” *Distributed Computing Systems Workshops*, pp. 575- 578, 2002.
- [10] F. L. Lewis. “Wireless sensor networks.” In D. J. Cook and S. K. Das, editors, *Smart Environments: Technologies, Protocols, and Applications*, New York, 2004. John Wiley.

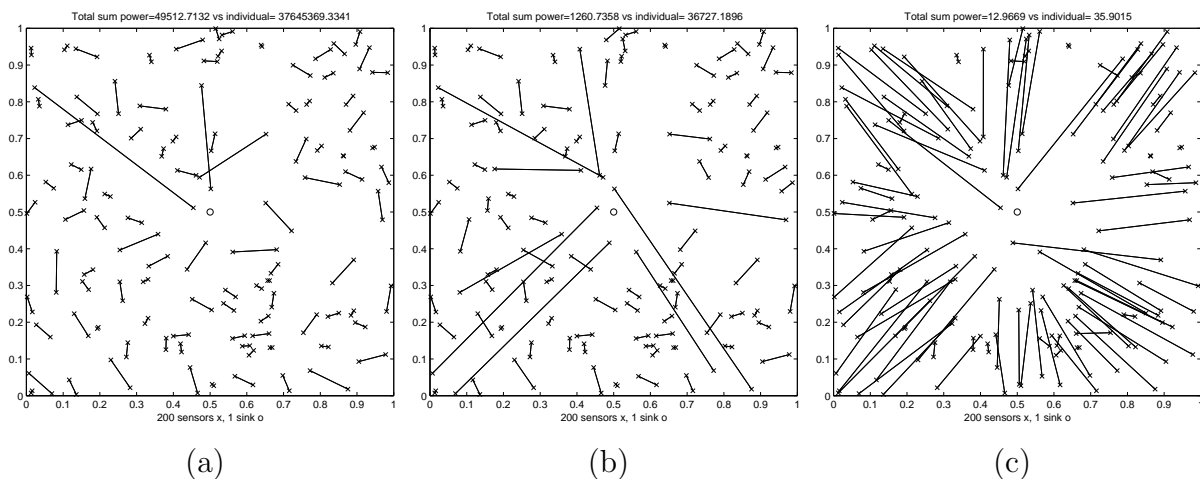


Figure 7: WSN with 200 sensors (x) and 1 sink (o) placed at the center. 2 matched nodes are linked by a line. Matching results for noisy channels without peak power constraint, and individual entropy is 20 in (a) 10 in (b), and 1 in (c).

- [11] C. H. Papadimitriou, and K. Steiglitz, *Combinatorial Optimization : Algorithms and Complexity*, Dover Pubns, July 1998.
- [12] S. Patten , B. Krishnamachari , R. Govindan, “The impact of spatial correlation on routing with compression in wireless sensor networks,” Proceedings of the third international symposium on Information processing in sensor networks, April 26-27, 2004.
- [13] A. Roumy, D. Gesbert, “Optimal matching in wireless sensor networks,” in Proc. IEEE International Symposium on Information Theory, ISIT 07, June 2007.
- [14] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information. sources,”. IEEE Trans. Inform. Theory., vol. IT-19, pp. 471-480, July 1973.
- [15] N. J. A. Sloane. Sequence A001147, in “The On-Line Encyclopedia of Integer Sequences.”
- [16] V. Stankovic, A. Liveris, Z. Xiong, and C. Georghiades, “On code design for the general Slepian-Wolf problem and for lossless multiterminal communication networks,” IEEE Trans. Inform. Theory, Vol. 52, Issue 4, pp. 1495-1507, April 2006.
- [17] V. Toto-Zarasoia, A. Roumy, and C. Guillemot, “Rate-adaptive codes for the entire Slepian Wolf region and arbitrarily correlated sources,” in preparation.

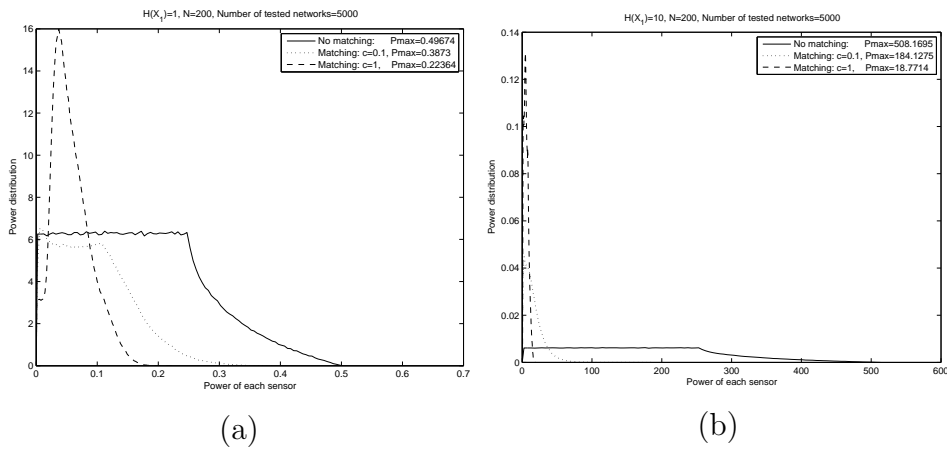


Figure 8: Noisy node-sink channel case. Distribution of the power $\{P_i\}_i$ consumed in the sensor network. Result of the sum power minimization, Problem 2, for different correlations c and different individual entropies $H(X_1)$. Curves obtained for 5000 realizations of a WSN with 200 sensors, sink placed at the center of the network. (a) $H(X_1) = 1$ (b) $H(X_1) = 10$.