Linear and Non-Linear Strategies for Power Mapping in Gaussian Sensor Networks

Franco Davoli, Mario Marchese, Maurizio Mongelli

DIST - Department of Communications, Computer and Systems Science
CNIT - National Inter-University Consortium for Telecommunications
University of Genoa, Via Opera Pia 13, 16145, Genova, Italy
{ franco.davoli, mario.marchese, maurizio.mongelli }@unige.it

Abstract—**This paper deals with non-linear coding-decoding strategies for Gaussian sensor networks that obey a global power constraint and are decentralized (each sensor's decision is based solely on the variable it observes). The sensors and the sink act as the members of a team, i.e., they possess different information and they share a common goal, which consists in minimizing the expected distortion on the variables of interest. As the inherent power allocation, derived in "static" conditions (stationarity of the stochastic environment, fixed topology), reveals to be optimal [1], the main interest is to analyze its robustness to variable system conditions. To this aim, this paper goes deep inside the generalization capabilities of the proposed approach, by showing some interesting insights into the structure of the problem. The overall surprising outcome is that a quasi-static application of the approach reveals to be sufficient to maintain suboptimal performance even under a dynamic environment.**

Keywords—Gaussian sensor networks, neural control, power allocation, sensitivity analysis

I. INTRODUCTION

THE DEPLOYMENT of sensor networks is often such that measurements acquired by the sensor nodes are conveyed toward a sink, where they need to be processed and analyzed. Recently, there has been a growing interest in understanding the peculiarities of wireless sensor networks from both an information theoretic and decision theoretic point of view, particularly when the measured quantities can be represented as Gaussian random variables (see, e.g., [1]-[3]). In particular, when the measured variables are analog quantities, they can either be transmitted as such, or they can be quantized and transmitted according to a digital scheme. Given a distortion measure for the reconstruction of the variables at the sink, and the statistical characteristics of the communication channel and of the sources, it is not straightforward to determine whether joint source-channel coding (with analog transmission) can outperform the separation that is typical of digital communications [2]. In [1], we introduced decentralized decision models in the setting of team theory [4], based on the approximation of the optimal decision strategies by means of fixed-structure parametrized nonlinear functions, by applying the Extended Ritz Method (ERIM) [5]. The reason for seeking numerical approximations to the optimal coding/decoding strategies stems from the fact that a team optimization approach to these problems presents formidable analytical difficulties (originally pointed out in [6], even in a scalar source-channel model).

More specifically, in [1] we used neural approximating functions to derive a non-uniform power distribution among the encoders at each sensor node. We show that, in the presence of correlated measurements from the same source representing a physical phenomenon, a large reduction in the overall transmission power can be obtained, at the expense of very little increase in distortion, with respect to linear encoding strategies. This is achieved by selecting a few "representative" sensors from the total available pool. The inherent optimization algorithm achieves optimal solutions under static conditions (stationarity of the stochastic environment, fixed topology). As such, there is the need of investigating:

- *how is the solution sensitive to changes in the topology and in the statistical environment*?
- *how much computational and bandwidth effort is required*?

A simulation-based sensitivity analysis of the approach is outlined here with respect to several system parameters. The rest of the paper is organized as follows. We define the problem formally in the next section. Section III outlines our functional approximation approach. The sensitivity analysis is presented in Section IV and conclusions in Section V.

II. PROBLEM STATEMENT

We consider a number N of sensors deployed over a geographical area, each one observing a realization of some physical phenomenon described by a random variable (r.v.) S (the source). We adopt the model of [3], which we describe in the following. We suppose the observations to take place at discrete time instants, but, since we are interested in real-time, single-letter coding, we do not introduce the time index in the following for simplicity of notation. Successive source outputs are uncorrelated; however, there is spatial correlation between the source and the event observed by sensor i, represented by the r.v. S_i. As a consequence, the r.v.'s S_i

and S_j are also mutually correlated. We indicate by $\rho_{s,i}$ and $\rho_{i,j}$ the correlation coefficients between S and S_i, and between S_i and S_j, respectively. Moreover, we suppose $S \sim \mathcal{N}\left(0,\sigma^2\right)$, and that all the other variables $S_1,...,S_N$ are jointly Gaussian, with 0 mean, the same variance σ^2, and covariance matrix Σ_S. Measurements are corrupted by observation noise, so that sensor *i* observes a realization of the r.v.

$$X_i = S_i + N_i \tag{1}$$

with $N_i \sim \mathcal{N}\left(0,\sigma_N^2\right), \forall i$. The measurements are encoded at each sensor according to some real-time coding strategy

$$Z_i = f_i\left(X_i\right) \tag{2}$$

and the sink receives a channel output of the type

$$\boldsymbol{Y} = col\left[Y_1...Y_N\right], Y_i = Z_i + W_i \tag{3}$$

with $W_i \sim \mathcal{N}\left(0,\sigma_W^2\right), \forall i$. As in [3] the W_i is ignored without loss of generality. The sink's decoding strategy is also real-time and given byA

$$\hat{S} = g\left(\boldsymbol{Y}\right) \tag{4}$$

Functions $f_i\left(\cdot\right), i=1,...,N$ and $g\left(\cdot\right)$ should be chosen to minimize the quadratic distortion measure

$$D = E\left\{\left(S-\hat{S}\right)^2\right\} \tag{5}$$

under the overall power constraint

$$\sum_{i=1}^{N} E\left\{Z_i^2\right\} \le \Gamma \tag{6}$$

This problem, which will be referred to as *Problem 1*, assumes the presence of multiple receiving antennas at the sink (i.e., of an additive Gaussian noise MIMO channel), characterized by an identity matrix. In the presence of a Gaussian vector channel with uniform variance of the noise components, linear (i.e., uncoded) strategies would be optimal, as noted in [7]; however, the optimum gain matrix would be non-diagonal and would therefore require a centralized solution, whereas we are seeking a decentralized one, where each sensor encodes its observed variable. In sensor networks, the focus is usually on optimizing network lifetime; here, we find out the best trade-off between energy consumption and distortion by activating specific nodes as long as they have enough power. After that, other nodes are selected under the same selection scheme. It is arguable that this approach should be better than rotating transmission responsibilities around many nodes (in the latter case, some energy is spent periodically for signaling). More details on this issue are however left open for future research.

A static covariance matrix describes the mutual correlation among the input of the sensors. It depends on the distance between each pair of source-sensor and sensor-sensor; we refer to it as *topological covariance matrix*.

In [3], the following coding strategies are adopted:

$$Z_i = f_i(X_i) = \sqrt{\frac{P_i}{\sigma^2 + \sigma_N^2}} \cdot X_i \tag{7}$$

where $P_i = \Gamma/N$ is the power limit of sensor *i*, and the sink decoding strategy is:

$$g(Y) = \frac{1}{N}\sum_{i=1}^{N} Y_i,\ Y_i = \frac{E\left[X_i Z_i\right]}{E\left[Z_i^2\right]} \cdot Z_i = \frac{\sigma^2}{\sigma^2 + \sigma_N^2} \cdot X_i \tag{8}$$

We will refer to (7) and (8) here as *linear strategies*. By exploiting the fact that the source observations are correlated, the minimum number of sensors that need to be activated to achieve nearly optimal distortion should be sought, out of the total number of deployed sensors. An example may help understand. We consider both source S and noise in (1) having standard normal distributions. Fig. 1 represents a possible deployment of 30 sensors over a 50x50 grid and some related candidate subsets of sensors to be turned on (obtained by a simulation campaign); each element of the topological covariance matrix, with indexes i, j, is given by $\sigma^2 \cdot e^{-d_{ij}/10}$, according to a power exponential covariance model, d_{ij} being the distance between nodes i and j. A method for the determination of the optimal subset of sensors, different from the one we derive here, is provided in reference [3]; an advantage of the neural strategies we derive and apply in the following consists in the simultaneous optimal power assignment to the selected sensors.

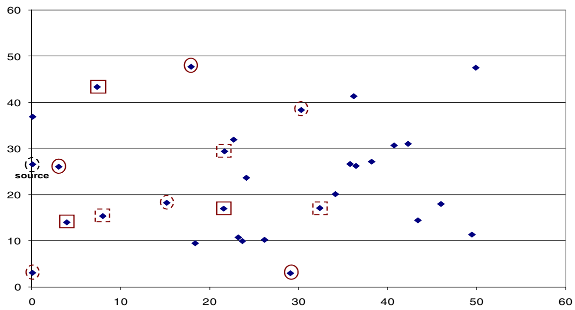

Figure 1. Example of deployment and candidate subsets to optimality.

III. Nonlinear Parametric Approximation of the Optimal Strategies

In this perspective, we reformulated Problem 1 in [1], by letting the coding and decoding strategies (2) and (4) depend on some non-linear approximation scheme. We remark again that the coding-decoding strategies are derived for a static topological covariance matrix. This assumption will be relaxed later. Introducing non-linear approximators in equations (2) and (4) means replacing them with:

$$Z_i = \hat{f}_i(X_i, \boldsymbol{w}_{f_i}) \tag{9}$$

$$\hat{S} = \hat{g}(\boldsymbol{Y}, \boldsymbol{w}_g) \quad (10)$$

where $\hat{f}(\cdot)$ and $\hat{g}(\cdot)$ are neural networks depending on the choice of the basis functions (e.g., sigmoidal) of each layer, and $\boldsymbol{w}_{f_i}$ and $\boldsymbol{w}_g$ are vectors of parameters activating the basis functions. Let $\boldsymbol{w}_f = col\left[\boldsymbol{w}_{f_1}, ..., \boldsymbol{w}_{f_N}\right]$. As already mentioned, the application of the entire vector $\boldsymbol{Y} = col\left[Z_1, ..., Z_N\right]$ in (10) helps highlight the performance gain induced by taking into account the topological structure (through explicit consideration of the cross-correlation in the strategies). Equations (9) and (10) are called *neural coding* and *decoding strategies*. Replacing (2) and (4) in the cost (5) with the neural strategies leads to the following parametric optimization problem (*Problem 2*):

$$\begin{gathered}\boldsymbol{w}_f^o, \boldsymbol{w}_g^o = \arg \min_{\boldsymbol{w}_f, \boldsymbol{w}_g} J(\boldsymbol{w}_f, \boldsymbol{w}_g);\ J(\boldsymbol{w}_f, \boldsymbol{w}_g) = E\left\{\left(S - \hat{S}\right)^2\right\}; \\ \hat{S} = \hat{g}\left(\left[\hat{f}_1(X_1, \boldsymbol{w}_{f_1}), ..., \hat{f}_N(X_N, \boldsymbol{w}_{f_N})\right], \boldsymbol{w}_g\right);\ i = 1, ..., N\end{gathered} \quad (11)$$

in order to find the optimal neural strategies $\hat{f}_i^o(\cdot) = \hat{f}_i(\cdot, \boldsymbol{w}_{f_i}^o)$ and $\hat{g}^o(\cdot) = \hat{g}(\cdot, \boldsymbol{w}_g^o)$ under the power constraint (6): $\boldsymbol{w}_{f_i},\ i = 1, ..., N : \sum_{i=1}^{N} E\left\{\hat{f}_i^2(X_i, \boldsymbol{w}_{f_i})\right\} \le \Gamma$.

Some technical details about the solution of Problem 2 can be found in reference [1]. Resorting from the team decision formulation of the functional optimization Problem 1 to the parametric approximation of Problem 2 is just an application of the methodology known as *Extended Ritz* [5]. Since a closed-form expression of the expected cost $J(\cdot)$ in (9) is not easily available, $J(\cdot)$ is substituted by its Montecarlo estimation $\tilde{J}(\cdot)$, $\tilde{J}(\cdot)$ being an arithmetic average over a given number Ξ of realizations of the random variables. More specifically, Ξ different samples of X_i and N_i, $i = 1, ..., N$, are generated on the basis of the topological covariance matrix and the distortion is computed under a given structure of the neural strategies (i.e., $\boldsymbol{w}_{f_i}$ and $\boldsymbol{w}_g$ are fixed). In doing this, the cost is also "augmented", by adding a penalty term, to ensure the satisfaction of the power constraint. The inherent numerical approximation is thus adaptive to general stochastic environments. The solution of (11) is obtained by applying an appropriate training chain of neural networks, which are placed inside each sensor and inside the sink. The training phase follows the regular back propagation algorithm for training neural network. In this perspective, the inherent computational effort is related to the off-line minimization of (11); on line, the trained neural networks can be applied "almost instantly". Disregarding the inherent application of a localization algorithm, our approach therefore needs a small on-line computational effort. Other aspects related to the cost of a localization system are discussed in subsection IV.E.

IV. Performance Analysis and Discussion

The analysis starts from the network of Fig. 1. The source S and noise in (1) have normal distributions (with unitary variances). We suppose that for linear strategies each sensor cannot transmit with more than one unit of power. The constraint over the overall power consumption used in Problem 2 is therefore $\Gamma = 30$. A C++ simulator was developed to for the mentioned neural training algorithm to solve (11) and to compute the inherent distortion and power performance. The training phase took 18 minutes over an Intel processor@1.73GHz.

A. Optimal power allocation

The power allocation at the end of training is exactly Γ. Despite all sensors are active, the final distortion after training is only 4% over the one guaranteed by linear strategies (*D*=0.95 for linear strategies versus *D*=0.99 for neural ones). The resulting power allocation, depicted in Fig. 2, states an important difference from linear strategies. The possible candidate subsets of sensors to be turned on are: {3, 8, 14, 18, 19} or {8, 18, 19}, whose power level is above or just in proximity of the line corresponding to *P*=1.5 in Fig. 2 (a discriminating power level for sensor activation).

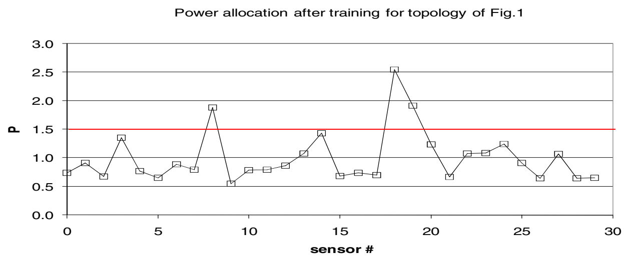

Figure 2. Power allocation after training (topology in Fig. 1).

	Distortion	Power
Linear Strategies with N=30	***0.952***	***30***
Linear Strategies {8,18,19}	1.151	3
Neural Strategies {8,18,19}	0.999	3.72
Linear Sensor - Neural Sink {8,18,19}	**0.999**	**3**
Linear Sensor - Optimal Linear Decoder {8,18,19}	1.015	3
Neural Sensor - Linear Sink {8,18,19}	1.653	3.73
Linear Strategies {3,8,14,18,19}	1.1	5
Neural Strategies {3,8,14,18,19}	0.999	5.79
Linear Sensor - Neural Sink {3,8,14,18,19}	0.999	5
Neural Sensor - Linear Sink {3,8,14,18,19}	1.52	5.8

Table 1. Performance under different combinations of strategies.

Table 1 reports the distortion and power consumption of different combinations of linear and neural strategies for the topology under investigation. The best subset reveals to be {8, 18, 19} (more marked circles in Fig. 3), because the other one only introduces a power waste. The most effective improvement is made by the introduction of the neural strategy at the sink, because using the linear strategies (for the discovered subsets), together with neural decoding, guarantees the best performance. This is also corroborated by a comparison with the optimal linear decoder for a Gaussian channel ('Linear Sensor - Optimal Linear Decoder {8,18,19}' in Table 1), in place of the neural sink, applied to the subset {8, 18, 19}. Such a decoder, jointly with the linear coding strategies, explicitly exploits the topological covariance matrix. The neural sink outperforms the optimal linear decoder because the latter is in any case a consequence of the simple linear coding scheme, which oversimplifies the structure of the linear decoder itself.

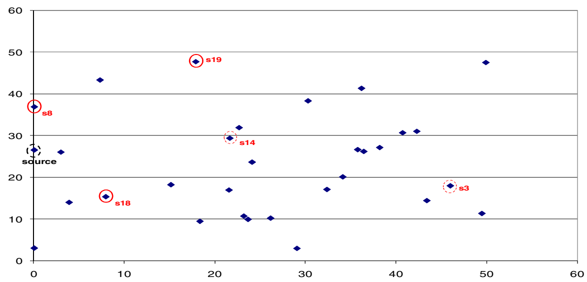

Figure 3. Power allocation after training (topology in Fig. 1).

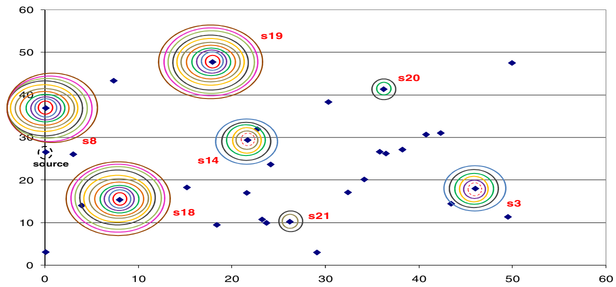

Figure 4. Activations using combinations of source and noise variances.

B. *Variable variances*

In order to validate the robustness of the neural power allocation, we consider the repetition of several training phases with respect to different combinations of source and noise variances. We consider again the topology of Fig. 1. Variances in play vary uniformly in the interval 5÷30, both for source and noise. A training phase is started from the beginning for each combination of variances. The qualitative result is depicted in Fig. 4, where a circle is marked around a sensor each time that sensor is chosen after training in virtue of its final power allocation (a simple comparison is made at the end of training as done for Fig. 2, on the basis of the automatic discriminating power level outlined above).

Fig. 4 means that the discovered optimal choice for sensor activation (for most of the times: sensors {8, 18, 19}) is invariant to the source and noise statistical behaviors. The rationale of this effect can be explained as follows. It is arguable that the final result in the power allocation is a function of the topological covariance matrix; in other words, the power solution is influenced by the reciprocal correlation of each couple of sensors and source-sensor. In this view, changing the specific statistical behavior of source and noise does not lead to significant changes in the sensor activation scheme. The neural approach "captures" the reciprocal influence coming from the spatial correlation coefficients and discovers an optimal power allocation scheme, which is thus robust to possible changes in the statistical environment. This has been validated by considering several other topologies. On the other hand, a radical change in the topology itself is more critical. The inherent effects can be explained as follows. A new training phase is needed when a significant topology change takes place. For instance, this means that the results in Fig. 4 significantly change if a topological permutation is applied over a "large" subsets of nodes. How the power allocation may be invariant to the "size" of such a permutation is an intriguing question as well. A method to match variable topologies jointly to the proposed neural scheme is outlined in the following.

C. *Different test sets*

Before that, one final remark is worth to be mentioned on the robustness of the neural scheme. It is related to the distortion obtained by the 'Linear Sensor Neural Sink {8,18,19}' technique (which is the best compromise between distortion and power consumption as to Table 1 above) in comparison with the 'Linear Strategies {8,18,19}', in the presence of variances different from the ones used during training. This means that we test the two techniques above with some 'test sets', which are different from the ones used for training. More specifically, these sets collect the sequences of samples (10^5, in our case) produced by the source and noise probability distributions during a simulation, which calculates the distortion (and the power allocation) via a Montecarlo approximation. The training set belongs to given normal distributions and the test sets to different normal distributions with variable variances. We denote by δ_d the quadratic difference between the distortion obtained at the end of training (i.e., over the set extracted during the last training step) and the one obtained with a specific test set.

Fig. 5 shows δ_d as a function of increasing source variance (σ^2) and noise variance (σ_N^2) used in the test sets ($\sigma^2 = \sigma_N^2 = 1$, over the training set). In the linear approach, the linear coding formula is used with $\sigma^2 = \sigma_N^2 = 1$ and an "ideal" distortion d^* is calculated using samples generated from normal distributions (coherently with the expected $\sigma^2 = \sigma_N^2 = 1$); δ_d is thus computed in the linear case as the quadratic difference between d^* and the distortion obtained using samples generated from Gaussian distributions with specific variances, namely, with the coding formula not updated with respect to the real variances used to generate the samples. From Fig. 5 it is quite evident that the neural approach outperforms the linear one as it limits the error

(δ_d) introduced by the application of test sets, whose samples are progressively different from the expected ones.

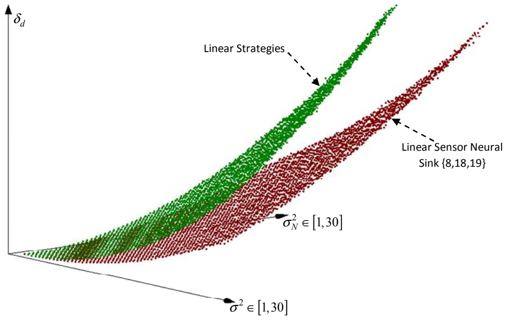

Figure 5. Variable variances: distortion error δ_d .

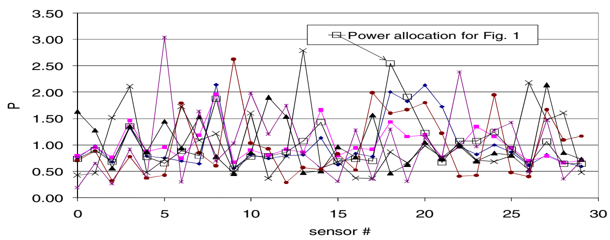

Figure 6. Activations using combinations of source and noise variances.

D. *Variable topologies*

Similar results have been obtained with other kinds of topologies and considering the presence of 100 sensors. Fig. 6 shows the result of our neural analysis for other topologies of 30 nodes (the other parameter settings of the topology in Fig. 1 are left untouched).

The neural strategies are capable to always discover the best subsets of sensors. To summarize, it is sufficient to discover the optimal subset and then, using some signaling scheme, activate the subset and turn off the other sensors; this action can be performed by some centralized unit, usually the sink node itself. Thus, one last and most crucial question naturally arises: *how can we discover the optimal subset with time-varying topologies without restarting the training phase from scratch?* In other words, the ultimate goal is to obtain a method capable to let the network learn the optimal subset of sensors to be turned on in dependence of the current topology. The method can be informally described as follows. First of all, a localization technique is needed for the sink to know the current geometry of the topology (information about the reciprocal distances among the nodes is sufficient). Then, the family of optimal power consumption curves (as in Fig. 9) can be learned off-line by the sink, by using again a neural approximation scheme. More specifically, since each curve in Fig. 6 derives from a specific topology (from a specific topological covariance matrix), a given function exists that maps the set composed of optimal power allocations with the set of corresponding topologies. The numerical approximation of this function, called optimal power mapping, can be easily obtained, because it consists of solving a problem easier than Problem 1 or 2 (actually, it does not involve any random quantity). A similar approach has been successfully employed in [8] to approximate the solutions of a (computationally expensive) pricing optimization problem, as a function of variable network bottlenecks and traffic demands. Once the optimal power mapping is updated at the sink, together with the current distances among the nodes, the sink becomes capable to optimize the performance (as discussed for Table 1) with a small computational effort. The prices to pay are a given amount of bandwidth dedicated to turn on and off the sensors and the adoption of a localization system. More specifically, the sink periodically updates the actual position of the sensors and computes the optimal power mapping if a change is needed in the subset of active sensors. As said in subsection IV.B, this mainly depends on the reciprocal correlation of each couple of nodes. In turn, it thus depends on the reciprocal distance among the nodes. No changes are needed if the topology of the network remains fixed. The approach is reasonable if the topology changes slowly with time; in this viewpoint, some time is needed for the localization system to recognize the possible movements of the nodes.

E. *Localization*

In this perspective, some additional words are necessary concerning localization. Several techniques are available in the literature concerning localization over sensor networks. As explained in the previous subsection, what is needed here is making available at the sink the set of the relative distances among the nodes. This does not involve the adoption of complicated localization approaches by using, for example, anchors-based schemes, which are suited to derive absolute coordinates in place of relative ones (see, e.g., chapter 9 in [9]). However, the problem of how much the accuracy of the localization system influences the optimal power mapping is the critical issue. As such, we analyze the sensitivity of the optimal power mapping as a function of the position of the source. The rationale behind this choice relies on the fact that source localization is the most critical issue due to source mobility (which is actually more likely to occur than the mobility of the sensors). Thus, without entering any detail, we must note that updating the optimal power mapping frequently by the localization system might result in a bandwidth and computational burden, which would make the application of the proposed system quite impractical. This is however not the case, in virtue of the sensitivity properties of the optimal power mapping with respect to topological changes. An example is reported in Fig. 7, where the optimal subsets of sensors are highlighted as a function of different positions of the source. The source is indicated by the different squares in Fig. 7.

Actually, the source moves around a cross centered in the middle of the sensing field (30 sensors are randomly deployed over a 50×50 square as in Fig. 1; source and noise have normal distributions). As done for Fig. 4, a circle is marked around a sensor each time that sensor is chosen after training in virtue of its final power allocation. Despite the different positions of the source, an optimal subset of sensors is derived by the neural analysis, which is thus robust to the actual position of the source. This is corroborated by Table 2, where the quadratic difference, δ_d, between d^* (the distortion obtained by the adoption of the optimal subsets obtained in correspondence of each new position of the source) and $d^{'}$ (the distortion achieved in all cases with the optimal subset obtained for the source placed in the center of the field) is outlined. The δ_d quantity is always very low (several orders of magnitude lower than the absolute distortion values, which are around 1.0), thus meaning that the optimal power mapping can be actually not aware of the exact location of the source, without incurring any distortion degradation. In other words, the optimal subset of sensors obtained when the source is in the center of the square reveals to be a suboptimal choice for all of the other cases when the source is located elsewhere. The term "suboptimal" here is related to the actual performance of our approach with an imperfect localization system that does not always achieve a perfect knowledge of the nodes' position.

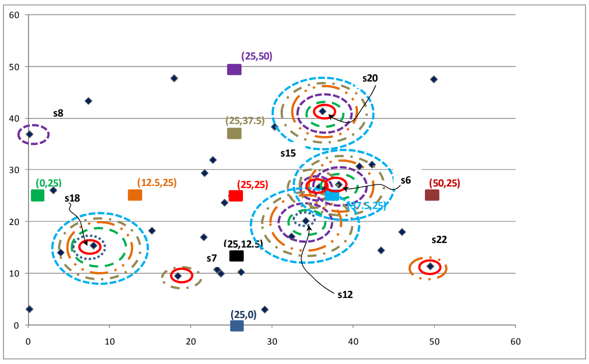

Figure 7. Activation with different source positions (squares in the picture).

Position of the source	Subsets after training	δ_d
25-25 (center of the square)	6,18,20	0
25-0	12,18	4.00E-04
0-25	12,18,20	1.00E-04
25-50	6,12,20	4.84E-04
50-25	12,18,20	4.72E-04
12,5-25	12,18,20	4.00E-04
25-12,5	12,18	1.44E-04
25-37,5	6,12,18,20	1.00E-06
37,5-25	12,18	1.44E-04

Table 2. Different optimal subsets and source positions: distortion error δ_d.

The generalization of this result leads to the conclusion that a localization method with stringent time constraints is not mandatory. The price to pay for updating localization data over large time scales (thus, with respect to the bandwidth and computational constraints of the actual technology in play) is obtaining suboptimal distortion performance, which arises when the optimal power mapping is not updated with the actual distances among the nodes. This is however acceptable, in virtue of the provided power saving. We verified similar results to Fig. 7 and Table 2 (almost distortion-invariance of the optimal power mapping to topology changes), by changing the location of a subset of sensors without updating coherently the optimal power mapping, until the cardinality of that subset becomes higher than 32.5% of the total number of nodes. Any increase over that threshold in the number of sensors, (whose locations the optimal power mapping is not aware of) implies distortion deterioration over 42.6% of the optimal one (this was validated through an extensive simulation campaign with up to 100 nodes). That threshold can be considered in an empirical rule of thumb to define when updating the optimal power mapping with actual localization data would be necessary. In practice, as having a sudden dislocation of more than 32.5% of the nodes can be considered a rare event (almost impossible in several applications where the positions of the sensors are fixed), at least suboptimal performance is always guaranteed.

V. Conclusions and Future Work

The paper has analyzed the performance sensitivity of the non-linear optimization approach to Gaussian sensor networks of [1]. The overall outcome is that a quasi-static application of the approach is sufficient to maintain suboptimal performance under a dynamic environment. Future work concerns exporting the analysis towards real scenarios beyond the Gaussian setting taken as a reference in this paper.

References

[1] F. Davoli, M. Marchese, M. Mongelli, "A Decision Theoretic Approach to Gaussian Sensor Networks," *Proc. IEEE Internat. Conf. Commun. 2009 (ICC 2009)*, Dresden, Germany, June 2009.

[2] M. Gastpar, M. Vetterli, P. L. Dragotti, "Sensing Reality and Communicating bits: A Dangerous Liaison," *IEEE Signal Processing Mag.*, vol. 23, no. 4, pp. 70-83, July 2006.

[3] M. C. Vuran, I. F. Akyildiz, "Spatial Correlation-Based Collaborative Medium Access Control in Wireless Sensor Networks," *IEEE Trans. Networking*, vol. 14, , no. 2, pp. 316–329, April 2006.

[4] Y. C. Ho, M. P. Kastner, E. Wong. "Teams, Signaling, and Information Theory," *IEEE Trans. Automat. Contr.*, vol. AC-23, pp. 305-311, 1987.

[5] R. Zoppoli, M. Sanguineti, T. Parisini, "Approximating Networks and Extended Ritz Method for the Solution of Functional Optimization Problems," *J. of Optimization Theory and Applic.*, vol. 112, no. 2, pp. 403-439, Feb. 2002.

[6] H. S. Witsenhausen. "A Counterexample in Stochastic Optimum Control," *SIAM Journal of Control*, vol. 6, pp. 131-147, 1968.

[7] R. J. Pilc, "The Optimum Linear Modulator for a Gaussian Source used with a Gaussian Channel", Bell Syst. Tech. J., vol. 48, pp. 3075-3089, Nov. 1969.

[8] F. Davoli, M. Marchese, M. Mongelli, "Neural Decision Making for Decentralized Pricing-based Call Admission Control," *Proc. IEEE Internat. Conf. Commun.* (*ICC 2005*), Seoul, Korea, May 2005.

[9] H. Karl, A. Willig, *Protocols and Architectures for Wireless Sensor Networks*, John Wiley and Sons Ltd., Chichester, 2005.