This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2009 proceedings

A Decision Theoretic Approach to Gaussian Sensor Networks

F. Davoli, M. Marchese, M. Mongelli

DIST - Department of Communications, Computer and Systems Science
University of Genoa, Via Opera Pia 13, 16145, Genova, Italy
{ franco, dama, mopa }@dist.unige.it

Abstract—**We consider the acquisition of measurements from a source, representing a physical phenomenon, by means of sensors deployed at different distances, and measuring random variables that are correlated with the source output. The acquired values are transmitted to a sink, where an estimation of the source has to be constructed, according to a given distortion criterion. In the presence of Gaussian random variables and a Gaussian vector channel, we are seeking optimum real-time joint source-channel encoder-decoder pairs that achieve a distortion sufficiently close to the theoretically optimal one, under a global power constraint, by activating only a subset of the sensors. The problem is posed in a team decision theoretic framework, and the optimal strategies are approximated by means of neural networks. We compare the solution with the results obtained by heuristically choosing a subset of the sensors on the basis of successive simulations under a fixed topology.**

Keywords—Gaussian sensor network, neural control

I. INTRODUCTION

SENSOR networks are often employed to measure some physical quantities of interest (e.g., temperature, pressure, concentration of chemicals, etc.), with the measurements being distributed in space and time. It is also often the case that a number of sensors are spread in random fashion over the area of interest, and have to forward their measured values over a noisy wireless channel to one or more sinks, which then perform an estimation of the desired quantity, based on some fidelity criterion. In many situations, the quantities being observed are analog ones, and can be represented by continuous random variables, as can the noise on the communication channels. Therefore, when looking at the transmission and estimation problem, one is confronted with the choice of whether to adopt digital transmission, and accordingly try to find encoder-decoder pairs that follow the separation theorem of information theory [1] (achieving optimality asymptotically in the transmission delay and disregarding complexity), or to remain in the domain of analog signals and seek joint source-channel coding pairs, with the additional requirement of being constrained in block length (and then in coding complexity and delay) [2, 3].

Indeed, besides the well-known case of a scalar Gaussian random variable to be transmitted over a Gaussian channel, where uncoded and "instantaneous" (also known as "single-letter" or "real-time") transmission is optimal (in the sense of achieving minimum quadratic distortion under a power constraint) [4], there are recent interesting results that prove the optimality of such uncoded schemes for a class of Gaussian sensor networks in general [5] (after a scaling-law optimality of uncoded transmission had been shown previously to hold for the same class of networks [6, 7]).

Notwithstanding these cases, however, the separation theorem would still apply in other more general situations as, among others, the Gaussian vector channel considered in [8] (where the best *linear* encoder-decoder pairs are sought). Interestingly enough, the best linear solution here (i.e., the uncoded one, in the sense that the input to the channel is only obtained by multiplying the observed values by a matrix, subject to the overall power constraint) turns out to be also instantaneous (i.e., with block length of 1), and may prevent some components with low signal-to-noise ratio to be transmitted, in favor of others. General conditions on the optimality of real-time (or "single-letter") encoders and decoders have been derived in [3], and the general structure of real-time encoding-decoding strategies for general Markov sources is analyzed in [9]. It is worth recalling, however, that even constraining the block length to the minimum, the optimal strategies that map the available information into coding and decoding decisions are in general unknown - with the exception of the special Gaussian cases cited above – and their determination entails the solution of a team optimization problem with dynamic information structure [10], which raises formidable difficulties.

Within this framework, we consider the problem of the acquisition of measurements from a Gaussian source, representing a physical phenomenon, by means of sensors deployed at different distances, which measure random variables that are correlated with the source output and corrupted by Gaussian noise. The acquired values are transmitted to a sink (under an overall power constraint), where an estimation of the source has to be constructed, according to a quadratic distortion criterion. The presence of correlated measurements intuitively suggests that a distortion value sufficiently close to the optimum one (under a given power constraint) might be achieved by activating only a subset of sensors among those deployed in the area, possibly at the price of a lower power consumption. This idea has been exploited in [11], where the authors determine the minimum number of observations from a subset of the whole population of sensors that is necessary to achieve a distortion very close to the optimal one (under the chosen encoding strategy), by means of successive simulative evaluations with a fixed topology. We are interested in investigating team decision strategies for the encoders and the decoder that solve the same problem. In other words, we want to deploy Decision Makers (DM) at the sensors and at the sink that are able to decide upon sensor activation and power assignments, in order to achieve the minimum distortion level within a given precision. Since the analytical determination of such team strategies is, in general, a formidable problem, we resort to nonlinear parametric approximation.

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. Numerical results are 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 [11], 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 $W_i \sim \mathcal{N}\left(0,\sigma_W^2\right)$, $\forall i$. The sink's decoding strategy is also real-time and given by

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

Functions $f_i(\cdot), i=1,...,N$ and $g(\cdot)$ 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. The observation channel would also exhibit an identity matrix (from eq. (1)); however, since we are observing r.v.'s characterized by a certain correlation matrix Σ_S, which depends on the sensors' deployment, we could change the observation channel (as noted in [7, footnote on p. 747]), by considering independent variables $\tilde{S}_1,...,\tilde{S}_N$, into one characterized by the matrix Q^H, where Q is the Karhunen-Loève transform of $S_1,...,S_N$, i.e., a unitary matrix such that $Q^H \Sigma_S Q$ is diagonal. We are therefore in a setting similar to that in [7], where uncoded transmission is shown to achieve asymptotic scaling law optimality in several cases, but not to be exactly optimal. It is worth noting that another interesting variant of the problem arises when transmission takes place over a Gaussian multiple access channel, where equation (3) would be substituted by

$$Y = \sum_{i=1}^{N} Z_i + W \tag{7}$$

With respect to encoding-decoding strategies adopted in [11], no difference exists between the two settings, as we will see shortly. We adopt the case where richer information is available, to better highlight the possible gains.

To simplify the analysis, the topology of the network is considered fixed, namely, no faults or movements of the sensors are possible. Thus, a static covariance matrix describes the mutual correlation among the input of the sensors. It depends on the distance between each pairs of source-sensor and sensor-sensor; we refer to it as *topological covariance matrix*. To operate in the same setting as [11] for comparison, the noise W_i in (3) is ignored from now on. Transmission noises could anyway be included straightforwardly in our treatment.

It is worth noting that, since the variance of the source components is uniform, 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 [8]; 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 [11], uncoded transmission is anyway adopted for the sensors; then, 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 is 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 ($\sigma^2 = \sigma_N^2 = 1$). Fig. 1 represents a possible deployment of 30 sensors over a 50x50 grid; 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, j. In [11], the following coding strategies are adopted:

$$Z_i = f_i(X_i) = \sqrt{\frac{P_i}{\sigma^2 + \sigma_N^2}} \cdot X_i \tag{8}$$

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

$$g(Y) = \frac{1}{N}\sum_{i=1}^{N} Y_i,\ Y_i = \frac{E\left[X_i Z_i\right]}{E\left[Z_i^2\right]} \cdot Z_i = \frac{\sigma^2}{\sigma^2 + \sigma_N^2} \cdot X_i \tag{9}$$

We will refer to (8) and (9) here as *linear strategies*.

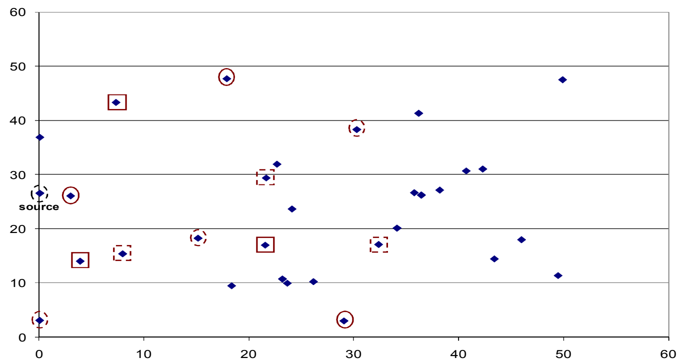

Figure 1. Example of deployment and candidate subsets to optimality.

We must note that, even if (8) and (9) result in non-exactly optimal strategies, they represent a reasonable choice for a distributed environment with possibly unknown sensor locations and a useful tool to derive insights into the performance sensitivity of Problem 1. No information about the spatial correlation is exploited by the linear strategies. This is a natural choice, due to the distributed solution of Problem 1 that is assumed in [11]. In situations where centralized strategies are applicable, optimality would be guaranteed by a centralized exploitation of the covariance matrix [8].

Turning back to the topology of Fig. 1, the distortion performance (5), obtained by linear strategies, is depicted in Fig. 2, by randomly choosing subsets of active sensors among the total available ones. On the x-axis we have different realizations of each subset (of n sensors out of N), and on the y-axis the corresponding distortion ($D(n)$). The power consumption is always equal to the number of active sensors (we assume each sensor spends a unit of power to transmit). Each point is the result of averaging the distortion over 10^5 samples of X_i and N_i, i=1,…,30. The performance decrease in distortion using 3 sensors, in place of 30, can be very low if the 3 sensors are accurately chosen, while the power revenue amounts to one order of magnitude. This result is qualitatively confirmed when changing the topology, as depicted in Fig. 3, where different topologies are randomly generated and the performance loss of $D(3)$ is outlined with respect to its minimum and average over 20 random samples of the subset within a given topology.

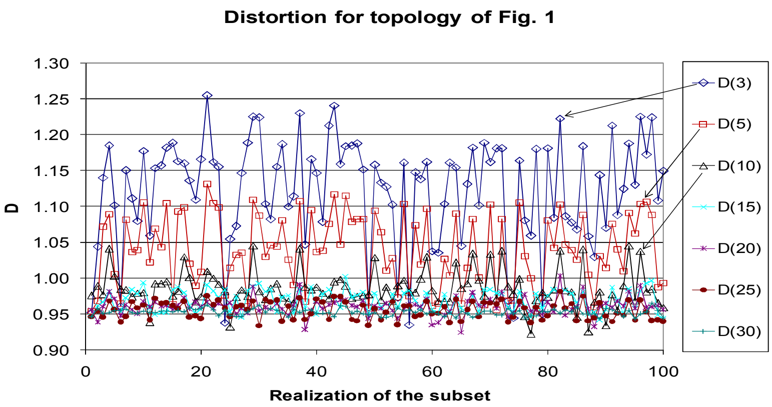

Figure 2. Distortion for the topology of Fig. 1.

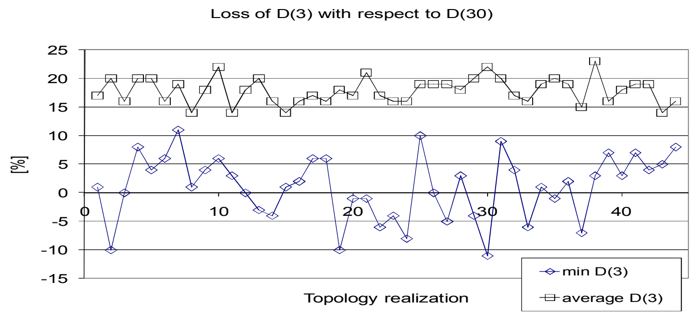

Figure 3. Loss in distortion using 3 sensors.

The performance loss is sometimes surprisingly negative, so that a specific subset is capable to achieve a distortion lower than $D(30)$ (under decoding strategy (9), which ignores the correlation). The final result is that applying a conservative power consumption scheme (by turning on only a small number of sensors) does not significantly decrease the distortion. Fig. 1 reports some candidate subsets to optimality under the rationale that one should activate no more than N/10 sensors, out of the N available, and they should be sufficiently far away from each other in order to minimize mutual interference [11].

The rationale under this analysis is that we know that the optimal subset depends on some geometric property of reciprocal sensors' positions, but how to let the network self-learn this subset is still an open issue.

To do that, we take a different approach here, by trying to determine the optimal strategies (2) and (4), and by then looking at the power distribution that they entail, which is not uniform among the sensors.

III. NONLINEAR PARAMETRIC APPROXIMATION OF THE OPTIMAL STRATEGIES

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

$$Z_i = \hat{f}_i(X_i, \boldsymbol{w}_{f_i}) \quad (10)$$

$$\hat{S} = \hat{g}(\boldsymbol{Y}, \boldsymbol{w}_g) \quad (11)$$

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 $\boldsymbol{w}_{\boldsymbol{f}} = col\left[\boldsymbol{w}_{f_1},...,\boldsymbol{w}_{f_N}\right]$. As already mentioned, the application of the entire vector $\boldsymbol{Y} = col\left[Z_1,...,Z_N\right]$ in (11), helps highlight the performance gain induced by taking into account the topological structure (through explicit consideration of the cross-correlation in the strategies). (10) and (11) 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*):

$$\begin{aligned}&\boldsymbol{w}_{\boldsymbol{f}}^o, \boldsymbol{w}_g^o = \arg\min_{\boldsymbol{w}_f, \boldsymbol{w}_g} J(\boldsymbol{w}_{\boldsymbol{f}}, \boldsymbol{w}_g);\ J(\boldsymbol{w}_{\boldsymbol{f}}, \boldsymbol{w}_g) = E\left\{\left(S-\hat{S}\right)^2\right\};\\ &\hat{S} = \hat{g}\left(\left[\hat{f}_1(X_1, \boldsymbol{w}_{f_1}),...,\hat{f}_N(X_N, \boldsymbol{w}_{f_N})\right], \boldsymbol{w}_g\right);\ i=1,...,N\end{aligned} \quad (12)$$

in order to find out 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): $\boldsymbol{w}_{f_i},\ i=1,...,N : \sum_{i=1}^{N} E\left\{\hat{f}_i^2(X_i, \boldsymbol{w}_{f_i})\right\} \le \Gamma$. How the optimal neural strategies are capable to introduce a performance gain in the distortion and to distribute the power among the sensors better than the linear strategies is studied in the next section.

Before that, some more words are necessary for the technical details about the solution of Problem 2. Resorting from a team decision formulation, lying on a functional optimization problem (Problem 1, in this case), to a parametric approximation (Problem 2) is known as *Extended Ritz* method [12]. Since a closed-form expression of the expected cost $J(\cdot)$ in (11) is not easily available, $J(\cdot)$ is substituted by its Montecarlo estimation $\tilde{J}(\cdot)$, being $\tilde{J}(\cdot)$ 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). Then, the next step (k+1) for $\boldsymbol{w}_{f_i}$ and $\boldsymbol{w}_g$ in the direction of the (hopefully global) minimum of $J(\cdot)$ is (for $i=1,...,N$):

$$\boldsymbol{w}_g(k+1) = \boldsymbol{w}_g(k) - \eta_g(k)\cdot\nabla_{\boldsymbol{w}_g}\tilde{J}^D(\boldsymbol{w}_{\boldsymbol{f}}(k), \boldsymbol{w}_g(k)) \quad (13)$$

$$\begin{aligned}&\boldsymbol{w}_{f_i}(k+1) =\\ &\boldsymbol{w}_{f_i}(k) - \eta_{f_i}(k)\cdot\left[\nabla_{\boldsymbol{w}_{f_i}}\tilde{J}^D(\boldsymbol{w}_{\boldsymbol{f}}(k), \boldsymbol{w}_g(k)) + \nabla_{\boldsymbol{w}_{f_i}}\tilde{J}^P(\boldsymbol{w}_{\boldsymbol{f}}(k), \boldsymbol{w}_g(k))\right]\end{aligned} \quad (14)$$

where $\tilde{J}^D(\cdot) = \left(S-\tilde{S}\right)^2$, $\tilde{S}$ being the output of the neural sink; $\tilde{J}^P(\cdot) = K_p\cdot\left(\sum_{i=1}^{N}\tilde{P}_i - \Gamma\right)^2$, $\tilde{P}_i$ being the square of the output of the i-th sensor. The quantities $\tilde{S}$ and $\tilde{P}_i$ are averaged over Ξ samples of X_i and N_i at each step k. Concerning the gradient of neural sensors, we know that their position in the decision chain is just before the sink; thus $\nabla_{\boldsymbol{w}_g}\tilde{J}^D(\cdot)$ and $\nabla_{\boldsymbol{w}_{f_i}}\tilde{J}^D(\cdot)$ are derived from the chain equations of the backpropagation algorithm used for training neural networks (not reported here for the sake of synthesis), initialized by $\nabla_{\boldsymbol{Y}}\tilde{J}^D(\cdot) = 2\left(S-\tilde{S}\right)$. $\nabla_{Z_i}\tilde{J}^P(\cdot)\Big|_{Z_i=\tilde{Z}_i} = 2K_P\cdot\left(\sum_{i=1}^{N}\tilde{P}_i - \Gamma\right)\cdot\tilde{Z}_i$, $\tilde{Z}_i$ being the output of the i-th sensor, averaged over Ξ samples of X_i and N_i at each step k. The application of the penalty cost function $\tilde{J}^P(\cdot)$ is necessary to match the power constraint (6). As far as convergence of (13) and (14) to the solution of Problem 2 is concerned, we must note that they belong to the family of *stochastic approximation* algorithms [13].

IV. PERFORMANCE EVALUATION AND DISCUSSION

The performance evaluation is related to the network of Fig. 1. 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 $\Gamma = 30$. Neural strategies are based on one-hidden layer neural networks with hyperbolic tangent neural units (one unit for each sensor and 5 units for the sink). Gradient stepsizes have both the form $1/(500+k)$; the penalty cost function parameter K_p is 0.25 and $\Xi = 10^5$. The training phase (13) and (14) took 18 minutes over an Intel processor@1.73GHz. The distortion during the training phase is depicted in Fig. 4. 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). Looking accurately at Fig. 2, one might argue that the neural strategies here fail optimizing the distortion performance, because two samples (out of 100) of D(3), in Fig. 2, can outperform D(30). This is probably due to the presence of local minima in the cost function of Problem 2 (a very frequent situation when training neural networks). However, this does not influence the most important information we derive from the neural analysis. The power allocation, depicted in Fig. 5, 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. 5 (a discriminating power level for sensor activation). Table 1 reports the distortion and power consumption of different combinations of linear and neural strategies for the topology under investigation. The best subset is {8, 18, 19} (more marked circles in Fig. 6), because the other one only introduces a power waste. Surprisingly, 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 means that linear strategies can be used without incurring in any performance decrease, but the neural sink helps limit distortion errors. It is also clear that neural strategies learn the principle of activating sensors close to the source and sufficiently separated to avoid reciprocal interference.

V. Conclusions and Future Work

The paper has presented a non-linear optimization approach to Gaussian sensor networks. The result obtained is the activation of a small number of sensors, while minimizing the power consumption without introducing significant loss in the distortion achieved at the sink. Future work may be exporting the analysis towards digital communications, by considering real scenarios beyond Gaussian hypotheses.

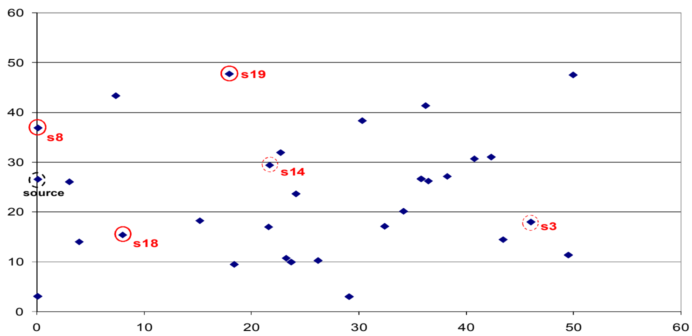

Figure 4. Power allocation after training (topology of Fig. 1).

	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**
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.

References

[1] C. E. Shannon, "A mathematical theory of communication", *Bell Sys. Tech. J.*, vol. 27, pp. 379–423 and pp. 623–656, July and Oct. 1948.

[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. Gastpar, B. Rimoldi, M. Vetterli, "To code, or not to code: Lossy source-channel communication revisited", *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1147-1158, May 2003.

[4] T. J. Goblick, "Theoretical limitations on the transmission of data from analog sources," *IEEE Trans. Inform. Theory*, vol. IT-11, pp. 558–567, Oct. 1965.

[5] M. Gastpar, "Uncoded transmission is exactly optimal for a simple Gaussian 'sensor' network", *Proc. 2007 Inform. Theory and Appl. Workshop*, San Diego, CA, Jan. 2007.

[6] M. Gastpar, M. Vetterli, "Source-channel communication in sensor networks", in F. Zhao, L. Guibas, Eds., IPSN 2003, *Lecture Notes in Computer Science* 2634, Springer-Verlag, Berlin-Heidelberg, 2003, pp. 162-177.

[7] M. Gastpar, M. Vetterli, "Power, spatio-temporal bandwidth, and distortion in large sensor networks", *IEEE J. Select. Areas Commun.*, vol. 23, no. 4, pp- 745-754, April 2005.

[8] 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.

[9] D. Teneketzis, "On the structure of optimal real-time encoders and decoders in noisy communication", *IEEE Trans. Inform. Theory*, vol. 52, no. 9, pp. 4017-4035, Sept. 2006.

[10] Y.-C. Ho, K.-C. Chu, "Team decision theory and information structures in optimal control problems – Part I", *IEEE Trans. Automat. Contr.*, vol. AC-17, no. 1, pp. 15-22, Feb. 1972.

[11] M. C. Vuran, Ö. B. Akan, I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks", *Computer Networks*, vol. 45, pp. 245–259, 2004.

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

[13] H. J. Kushner, G. G. Yin, *Stochastic Approximation Algorithms and Applications*, New York, NY, Springer-Verlag, 1997.

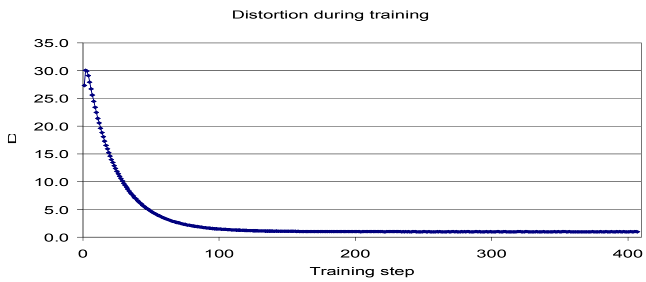

Figure 5. Distortion during training.

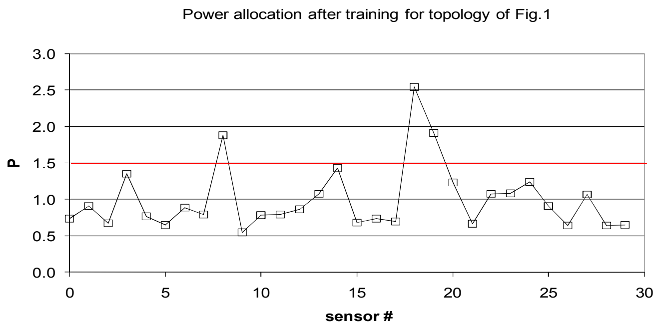

Figure 6. Power allocation after training (for topology of Fig. 1).