JOINT PRE-CODER DESIGN AND GREEDY POWER ALLOCATION FOR COMPRESSED SPATIAL FIELD ESTIMATION

Javier Matamoros and Carles Antón-Haro

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Av. Carl Friedrich Gauss 7, 08860 Castelldefels (Barcelona), Spain

ABSTRACT

In this paper, we propose a distributed beamforming scheme for the estimation of spatial fields (e.g. temperature, moisture) with wireless sensor networks. The pre-coding scheme allows for an over-the-air compressed representation of the correlated set of spatial observations, which are encoded in a number of consecutive sensor-to-gateway (GW) transmissions. The ultimate goal is to minimize the distortion in the reconstructed spatial field and, simultaneously, keep the number of transmissions low (i.e. the compression ratio high). However, the design of the set of normalized pre-coders, for which we derive a closed-form expression, and the corresponding power allocation problems turn out to be coupled. By resorting to a greedy power allocation strategy, both problems can be iteratively and jointly solved. The performance of the proposed pre-coding scheme is assessed by means of computer simulations. Other compressed beamforming schemes requiring channel inversion are used as a benchmark.

1. INTRODUCTION

In recent years, we have witnessed the emergence of the paradigm of Machine-to-Machine (M2M) communications [1].The M2M technical committee of ETSI (European Telecommunication Standards Institute) has proposed a hybrid architecture whereby cellular-enabled gateways (GW) act as traffic aggregation and protocol translation points for their capillary networks, typically based on short-range communication technologies (e.g. sensor networks). This paper focuses on the optimal design of such capillary extensions for environmental monitoring applications.

Our goal is to accurately reconstruct a spatial field from the samples collected by a number of sensing devices. Gastpar *et al* proved in [2] that cooperative *beamforming* turns out to be optimal when sensors intend to convey a *common* message (observation) to a remote destination. Unfortunately, this assumption does not hold here since our interest lies in monitoring the spatial *variations* of the field. A straightforward (yet not very efficient) approach would be to disseminate each observation to the rest of nodes prior to the beamforming stage. The inefficiency lies in the signalling overhead that such exchanges entail [3] and, also, in the fact that part of the exchanged information is known by the recipients (due to correlation). To circumvent that, in scenarios where signals are sparse, one can resort to compressed sensing techniques [4]. The cooperative beamforming approach adopted here is, thus, in stark contrast with the Amplify-and-Forward scheme of [5] where sensor observations are transmitted over *orthogonal* channels with no compression strategy in place. As for the fact that accurate phase synchronization over sensors is needed for distributed beamforming, the interested reader is referred to the synchronization strategy presented in [6].

In this paper, we propose an iterative greedy scheme allowing us to simultaneously solve the distributed beamforming (pre-coder) design and power allocation problems, which are inter-twined. By doing so, we go one step beyond our work in [**?**] which requires per-sensor channel equalization prior to the compressed beamforming phase. For the particular case of Gaussian channels, the iterative algorithm turns out to find the optimal solution, for which we also derive a closed-form expression.

This work is supported by the EXALTED (ICT-258512) and JUNTOS (TEC2010-17816) projects, and the Generalitat de Catalunya (2009 SGR 1046).

2. SIGNAL AND COMMUNICATION MODEL

Let $X(\mathbf{s})$ be a spatial field defined over the two-dimensional space $\mathbb{R}^2$. We assume that $X(\mathbf{s})$ is stationary, zero-mean and Gaussian-distributed. The spatial field is sampled by a set of N sensors located at $\mathbf{s}_1, \ldots, \mathbf{s}_N$ (locations are assumed to be known), this yielding

$$x_j \triangleq X(\mathbf{s}_j) \quad ; \quad j = 1, \ldots, N. \tag{1}$$

Consequently, the vector of observations $\mathbf{x} = [x_1, \ldots, x_N]^T$, where the variance of each component is σ_x^2, is jointly Gaussian and zero-mean too. For a specific set of locations, the elements of the covariance matrix $\mathbf{C}_\mathbf{x} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right]$ read $[\mathbf{C}_\mathbf{x}]_{j,j'} = \mathsf{k}(\mathbf{s}_j, \mathbf{s}_{j'})$, where $\mathsf{k}(\cdot,\cdot)$ denotes the covariance function of the spatial field. In addition, we let $\{\lambda_1, \lambda_2, \ldots, \lambda_N\}$ denote the eigenvalues of $\mathbf{C}_\mathbf{x}$ (without loss of generality, we assume $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$), and $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_N\}$ the corresponding eigenvectors.

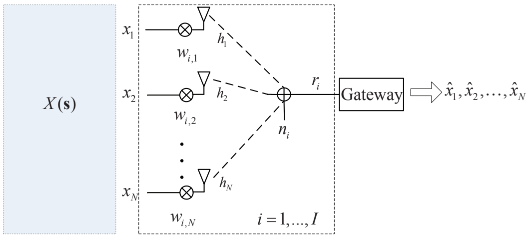

Fig. 1. Signal and communication model.

As shown in Fig. 1, sensors *simultaneously* transmit (i.e. beamform) their observations to the GW . For the i-th transmission, the received signal r_i reads

$$r_i \quad = \quad \sum_{j=1}^{N} w_{i,j} h_j x_j + n_i = \mathbf{w}_i^H \mathbf{H}\mathbf{x} + n_i \qquad (2)$$

for $i = 1, \ldots, I$, where $\mathbf{w}_i = \left[w_{i,1}, w_{i,2}, \ldots, w_{i,N}\right]^T$ denotes the pre-coder[1] (to be designed), the diagonal matrix $\mathbf{H} = \mathsf{diag}\left[h_1, h_2 \ldots, h_N\right]$ gathers the (complex) sensor-to-GW channel coefficients; and n_i is additive white Gaussian noise of variance σ_n^2, that is, $n_i \sim \mathcal{CN}\left(0, \sigma_n^2\right)$. Further, we assume slow fading conditions and, hence, the channel coefficients remain unchanged for the I consecutive transmissions. From the $I \times 1$ received vector $\mathbf{r} = \left[r_1, \ldots, r_I\right]^T$, the GW attempts to estimate (reconstruct) the spatial field at the set of sampled locations, namely, $\hat{\mathbf{x}}^{(I)} = [\hat{x}_1^{(I)}, \ldots, \hat{x}_N^{(I)}]^T$ where, for notational convenience, we make it explicit the dependency of the estimates on the total number of transmissions I. In the sequel, we assume $I \leq N$ and, hence, $\mathbf{r}$ can be regarded as a *compressed* representation of the observations vector $\mathbf{x}$. Due to channel impairments, noise and compression, the resulting estimates are subject to some distortion which will be characterized by the following quadratic metric:

$$D^{(I)} \triangleq \frac{1}{N} \sum_{j=1}^{N} \mathbb{E}\left[\left|\hat{x}_j^{(I)} - x_j\right|^2\right]. \qquad (3)$$

3. COMPRESSED TRANSMISSION

Our goal here is to find the set of pre-coders $\{\mathbf{w}_1, \ldots, \mathbf{w}_I\}$ and the associated transmit powers $\boldsymbol{\rho} = \{\rho_1, \ldots, \rho_I\}$ which minimize the distortion in the reconstructed spatial field. To start with, let $\mathbf{r}_{1:i-1} = \left[r_1, r_2, \ldots, r_{i-1}\right]$ denote the vector with the first $i-1$ elements (transmissions) in $\mathbf{r}$. From $\mathbf{r}_{1:i-1}$, the GW provides an MMSE estimate of the observations vector which is given by the posterior mean, namely,

$$\begin{aligned} \hat{\mathbf{x}}^{(i-1)} &= \mathbb{E}\left\{\mathbf{x}|r_1, r_2, \ldots, r_{i-1}\right\} \\ &= \mathbf{C}_{\mathbf{x}\mathbf{r}_{1:i-1}} \mathbf{C}_{\mathbf{r}_{1:i-1}}^{-1} \mathbf{r}_{1:i-1}, \end{aligned} \qquad (4)$$

where, in the above expressions, we have introduced the shorthand notation $\mathbf{C}_{\mathbf{x}\mathbf{r}} = \mathbb{E}\left[\mathbf{x}\mathbf{r}^H\right]$, and $\mathbf{C}_{\mathbf{r}} = \mathbb{E}\left[\mathbf{r}\mathbf{r}^H\right]$ to denote the corresponding covariance (subscripts have been omitted for brevity). The normalized average distortion after the $(i-1)$-th transmission thus reads:

$$D^{(i-1)} \quad = \quad \frac{1}{N} \mathrm{Tr}\left(\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}\right) \qquad (5)$$

where $\mathrm{Tr}\left(\cdot\right)$ denotes and $\mathbf{C}_{\mathbf{x}|\mathbf{r}} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T|\mathbf{r}\right]$ the posterior covariance matrix. By using the i-th transmission (i.e. increasing the number of transmissions by one), the current estimate of the spatial field can be successively refined, namely,

$$\begin{aligned} \hat{\mathbf{x}}^{(i)} &= \mathbb{E}\left\{\mathbf{x}|r_1, r_2, \ldots, r_{i-1}, r_i\right\} && (6) \\ &= \mathbf{C}_{\mathbf{x}\mathbf{r}_{1:i}} \mathbf{C}_{\mathbf{r}_{1:i}}^{-1} \mathbf{r}_{1:i} && (7) \end{aligned}$$

Since $\mathbf{x}$ and $\mathbf{r}_{1:i}$ are jointly Gaussian, the following identity holds

$$\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i}} = \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} - \frac{\mathbb{E}\left[\mathbf{x}r_i^*|\mathbf{r}_{1:i-1}\right] \mathbb{E}\left[r_i\mathbf{x}^T|\mathbf{r}_{1:i-1}\right]}{\mathbb{E}\left[r_i r_i^*|\mathbf{r}_{1:i-1}\right]} \qquad (8)$$

where[2] $\mathbb{E}\left[\mathbf{x}r_i^*|\mathbf{r}_{1:i-1}\right] = \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i$ and

$$\mathbb{E}\left[r_i r_i^*|\mathbf{r}_{1:i-1}\right] = \mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i + \sigma_n^2. \qquad (9)$$

From (8) again, the distortion after the i-th transmission, $D^{(i)} = \frac{1}{N}\mathrm{Tr}\left(\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i}}\right)$, can be recursively expressed as:

$$\begin{aligned} D^{(i)} &= D^{(i-1)} - \frac{1}{N} \mathrm{Tr}\left(\frac{\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i \mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}}{\mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i + \sigma_n^2}\right) \\ &= D^{(i-1)} - \frac{1}{N} \frac{\mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}^2 \mathbf{H}^H \mathbf{w}_i}{\mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i + \sigma_n^2}. \end{aligned} \qquad (10)$$

This allows us to find the i-th precoding vector such that it *successively* (and optimally) *refines* the previous estimate of the spatial field. In other words, the one which results into the lowest possible distortion $D^{(i)}$ given $D^{(i-1)}$.

3.1. Optimal pre-coders

From (10), the i-th pre-coding vector is given by the solution to the following optimization problem:

$$\max_{\mathbf{w}_i} \quad \frac{\mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}^2 \mathbf{H}^H \mathbf{w}_i}{\mathbf{w}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H \mathbf{w}_i + \sigma_n^2} \quad \text{s.to} \quad \|\mathbf{w}_i\|_2^2 \leq \frac{\rho_i}{\sigma_x^2}$$

with ρ_i denoting the power allocated to the i-th transmission:

$$\sum_{j=1}^{N} \mathbb{E}\left\{\left|w_{i,j}^* x_j\right|^2\right\} \quad = \quad \sigma_x^2 \|\mathbf{w}_i\|_2^2 \leq \rho_i. \qquad (11)$$

Clearly, the optimal solution will satisfy the above power constraint with equality and, thus, the optimization problem can be re-written as

$$\begin{aligned} &\max_{\tilde{\mathbf{w}}_i} && \frac{\tilde{\mathbf{w}}_i^H \mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}^2 \mathbf{H}^H \tilde{\mathbf{w}}_i}{\tilde{\mathbf{w}}_i^H \left(\mathbf{H} \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}} \mathbf{H}^H + \frac{\sigma_n^2 \sigma_x^2}{\rho_i} \mathbf{I}_N\right) \tilde{\mathbf{w}}_i} \\ &\text{s.to} && \|\tilde{\mathbf{w}}_i\|_2^2 = 1 \end{aligned}$$

[1]Notice that a different pre-coder is used for each transmission.

[2]For $i = 1$, the term $\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}$ in (8) must be replaced by $\mathbf{C}_{\mathbf{x}}$.

where $\tilde{\mathbf{w}}_i \triangleq \sqrt{\frac{\sigma_x^2}{\rho_i}}\mathbf{w}_i$ is the *normalized* pre-coder and $\mathbf{I}_N$ stands for the identity matrix of size N. Hence, the optimal normalized pre-coder is given by

$$\tilde{\mathbf{w}}_i^* = \lambda_{\mathsf{max}}\left\{\mathbf{H}\mathbf{C}^2_{\mathbf{x}|\mathbf{r}_{1:i-1}}\mathbf{H}^H, \mathbf{H}\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}\mathbf{H}^H + \frac{\sigma_n^2\sigma_x^2}{\rho_i}\mathbf{I}_N\right\} \tag{12}$$

where $\lambda_{\mathsf{max}}\{\mathbf{A},\mathbf{B}\}$ stands for the generalized eigenvector associated to the largest generalized eigenvalue of matrices $\mathbf{A}$ and $\mathbf{B}$. This last expression reveals that the pre-coder design and power allocation problems (to be addressed in the next subsection) are inter-twined: $\tilde{\mathbf{w}}_i^*$ depends not only on the transmit power allocated to the i-th transmission (through ρ_i) but, also, on the power allocated to all previous transmissions (through $\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:i-1}}$).

3.2. Optimal power allocation

The optimal power allocation strategy $\boldsymbol{\rho} = \{\rho_1, \ldots, \rho_I\}$ can be found by solving

$$\min_{\rho_1,\ldots,\rho_I,I} \quad D^{(I)} \qquad \text{s.to} \quad \sum_{i=1}^{I}\rho_i = P_t$$

with P_t denoting the total transmit power. It is worth noting that the minimization is over the set of transmit powers $\{\rho_i\}_{i=1}^{I}$ *and* the number of transmissions I. This, along with the coupling of the pre-coder design and power allocation problems, renders the problem not solvable analytically for the general case. However, a closed form solution exists for Gaussian channels, as the next section illustrates.

4. PARTICULAR CASE: GAUSSIAN CHANNELS

A closed-form solution will be found in two steps. First, we propose an iterative (and greedy) algorithm. Not only shall we realize that this approach is optimal for Gaussian channels but, also, the insights gained will allow us to propose an extension (and some justification) for the general case addressed in Section 5.

4.1. Iterative algorithm

For Gaussian channels, we have $\mathbf{H} = \mathbf{I}_N$ and, thus, equation (12) can be re-written as[3]

$$\begin{aligned}\tilde{\mathbf{w}}_l &= \lambda_{\mathsf{max}}\left\{\mathbf{C}^2_{\mathbf{x}|\mathbf{r}_{1:l-1}}, \mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:l-1}} + \frac{\sigma_n^2\sigma_x^2}{\rho_l}\mathbf{I}_N\right\} \quad (13)\\ &= \lambda_{\mathsf{max}}\left\{\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:l-1}}\right\} \quad (14)\end{aligned}$$

where the second equality follows from elementary properties of matrix algebra. Unlike in the general case, the design of the normalized pre-coder $\tilde{\mathbf{w}}_l^*$ here is no longer coupled with the power to be allocated to the l-th transmission itself. This

[3]For notational convenience, the transmission index i is replaced here by the iteration index l (see next paragraphs).

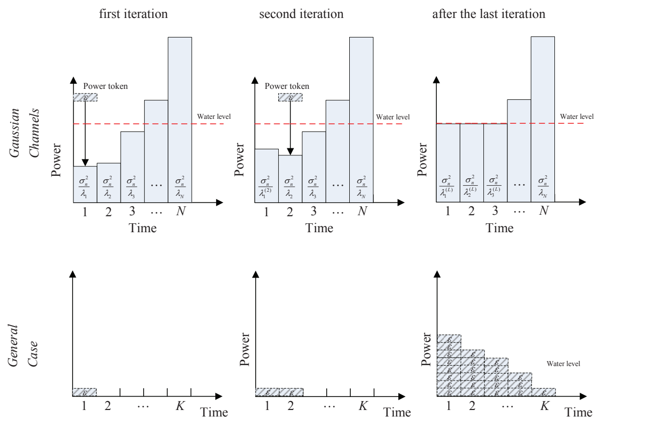

Fig. 2. Graphical representation of the greedy iterative power allocation scheme (Gaussian case)

considerably simplifies the problem at hand. In order to simultaneously solve the pre-coder design and power allocation problems, we propose to *iteratively* allocate transmit power in a greedy manner. To that aim, we define a *power token* ϵ as an indivisible and (sufficiently) small fraction of the total transmit power, namely $\epsilon \triangleq P_T/L$, where $L \gg 1$ stands for the total number of power tokens or iterations. For the first iteration ($l = 1$), it follows from (14) that $\tilde{\mathbf{w}}_1^* = \boldsymbol{\phi}_1$, that is, the eigenvector associated to λ_1, the largest eigenvalue of $\mathbf{C}_\mathbf{x}$. From equations (8)-(9) and since $\tilde{\mathbf{w}}_1^* = \boldsymbol{\phi}_1$, it follows that

$$\mathbf{C}_{\mathbf{x}|r_1} = \mathbf{C}_\mathbf{x} - \frac{\rho_1\lambda_1^2}{\rho_1\lambda_1 + \sigma_n^2\sigma_x^2}\boldsymbol{\phi}_1\boldsymbol{\phi}_1^H \tag{15}$$

and, hence, the eigenvectors of matrices $\mathbf{C}_{\mathbf{x}|r_1}$ and $\mathbf{C}_\mathbf{x}$ are identical. Clearly, this also applies to all matrices $\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:l}}$ to be drawn in subsequent iterations (but not for the *general* case, as will be discussed later). Since $\rho_1 = \epsilon$, from (15) we have that the eigenvalues of $\mathbf{C}_{\mathbf{x}|r_1}$, denoted by $\lambda_1^{(1)}, \lambda_2^{(1)}, \ldots, \lambda_N^{(1)}$ verify

$$\lambda_1^{(1)} = \lambda_1 - \frac{\epsilon\lambda_1^2}{\epsilon\lambda_1 + \sigma_n^2\sigma_x^2} \tag{16}$$

whereas $\lambda_k^{(1)} = \lambda_k$ for all $k \neq 1$. The power token in the second iteration will be allocated to the eigenvector associated to the largest eigenvalue out of $\lambda_1^{(1)} \ldots \lambda_N^{(1)}$. This iterative procedure is illustrated in Fig. 2. Note that, from (16), $\lambda_1^{(1)}$ is not necessarily the largest eigenvalue of $\mathbf{C}_{\mathbf{x}|r_1}$. In this case, the power token for the second transmission goes to a so far inactive eigenvector/eigenmode (e.g. $\lambda_2^{(1)}$ in Fig. 2). Otherwise, if $\lambda_1^{(2)}$ continues to be the largest eigenvalue, we have $\tilde{\mathbf{w}}_2^* = \boldsymbol{\phi}_1$ again. Accordingly, it can be proved that the eigenvalue of the resulting covariance matrix $\mathbf{C}_{\mathbf{x}|r_1,r_2}$ denoted by $\lambda_1^{(1)}$ reads

$$\lambda_1^{(2)} = \lambda_1 - \frac{(\rho_1+\rho_2)\lambda_1^2}{(\rho_1+\rho_2)\lambda_1 + \sigma_n^2\sigma_x^2} \tag{17}$$

Clearly, this is equivalent to allocate a power of $\rho_1 + \rho_2 = 2\epsilon$ and transmit just once with $\tilde{\mathbf{w}}_1^* = \phi_1$. After L iterations, and since the optimal pre-coders $\tilde{\mathbf{w}}_l^*$ to be used in *any* transmission necessarily belong to the set of N eigenvectors of the *unconditional* covariance matrix $\mathbf{C}_\mathbf{x}$, this iterative scheme leads to the waterfilling (and, thus, optimal) solution of the rightmost plot in Figure 2. This holds true as long as the power tokens are *small* enough since this allows all the eigenmodes to accurately reach the *common* waterlevel.

4.2. Equivalent closed-form solution

From all the above, the optimization problem (13) can be rewritten as

$$\min_{\rho_1,\ldots,\rho_N} \quad \sigma_x^2 - \frac{1}{N}\sum_{i=1}^{N} \frac{\rho_i \lambda_i^2}{\rho_i\lambda_i + \sigma_n^2\sigma_x^2} \tag{18}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \rho_i \leq P_t, \tag{19}$$

which (i) entails a minimization of the score function $D^{(N)}$ on $\{\rho_i\}_{i=1}^N$ only (N is fixed now); and (ii) is convex. The corresponding waterfilling-like solution is given by

$$\rho_i^* = \left[\frac{\sigma_n}{\sqrt{\mu}} - \frac{\sigma_n^2\sigma_x^2}{\lambda_i}\right]^+ \quad ; \quad i = 1, \ldots, N. \tag{20}$$

where $[x]^+ \triangleq \max\{x, 0\}$ and μ denotes the Lagrange multiplier associated to the power constraint, which can be computed as follows:

$$\mu = \left(\frac{\frac{P_t}{\sigma_x^2} + \sum_{i=1}^{K}\frac{\sigma_n^2\sigma_x^2}{\lambda_i}}{K\sigma_n}\right)^{-2}. \tag{21}$$

In this last expression, K stands for the largest number of transmissions such that (i) the optimal scaling factors verify $\rho_i^* = \frac{\sigma_w}{\sqrt{\mu}} - \frac{\sigma_n^2}{\lambda_i} \geq 0$ for $i = 1, \ldots, K$; and (ii) the sum-power constraint holds with equality, i.e. $\sum_{i=1}^{K} \rho_i^* = P_t$. In other words, the optimal number of transmissions is given by $I^* = K$ and, necessarily, $I^* \leq N$ (i.e. attains some compression).

5. GENERAL CASE: ARBITRARY CHANNELS

The iterative greedy algorithm to be presented here is largely inspired in that of Section 4.1. However, the fact that $\mathbf{H}$ is no longer an identity matrix has a substantial impact on the optimization problem. More precisely,

1. There is no straightforward relation between the solution to the generalized eigenvalue problem in (12) for different values of the transmission index i (or iteration index l). Here, neither eigenvectors are identical, nor only one of the eigenvalues changes through consecutive iterations. Essentially, all of them must be recomputed anew.
2. The design of the normalized pre-coder for the l-th iteration does depend on its own power token (and preceding ones too).
3. The problem requires an *explicit* optimization on I since it does not follow from the iterative power allocation or waterfilling scheme. As far as this paper is concerned, we resort to an exhaustive search over I.
4. The optimal number of transmissions I^* can (potentially) be larger than N since it is not upper bounded by the total number of different eigenvectors of $\mathbf{C}_x$.

All this, in turn, calls for a number of adaptations in the iterative scheme. As in the Gaussian case, however, the transmit power is allocated to the set of pre-coders on a token by token basis. In addition, no changes of previously allocated tokens are allowed (not an exhaustive search). For the sake of clarity, we introduce the shorthand notation $\boldsymbol{\phi}_1^{l-1}(\rho)$ to denote the eigenvector associated to the largest eigenvalue of the generalized eigenvalue problem in (12). The superscript $l-1$ accounts for the number of conditioning elements in the covariance matrix $\mathbf{C}_{\mathbf{x}|\mathbf{r}_{1:l-1}}$ in (12), while ρ is the *accumulated* power allocated to such eigenvector (including the current iteration). So, we fix the number of transmissions I and describe herinafter the iterative scheme for the $I = 3$ case:

First iteration ($l = 1$): The first power token ϵ is necessarily allocated to $\boldsymbol{\phi}_1^o(\epsilon)$. It is retained as the best pre-coder/power allocation combination so far and, hence, will be part of all the combinations in subsequent iterations.

Second iteration ($l = 2$): The allocation of the new power token results into two possible combinations of pre-coders and powers (i) $\{\boldsymbol{\phi}_1^o(\epsilon+\epsilon)\}$, one transmission (pre-coder); or (ii) $\left\{\boldsymbol{\phi}_1^o(\epsilon), \boldsymbol{\phi}_1^1(\epsilon)\right\}$, two transmissions. The resulting distortion is then computed for both combinations according to (10). Assume that (ii) attains the lowest distortion so far and, thus, this combination is retained.

Third iteration ($l = 3$): There exist three possible combinations for the allocation of the new power token, namely, (i) $\left\{\boldsymbol{\phi}_1^o(\epsilon+\epsilon), \boldsymbol{\phi}_1^1(\epsilon)\right\}$, with two transmissions; or (ii) $\left\{\boldsymbol{\phi}_1^o(\epsilon), \boldsymbol{\phi}_1^1(\epsilon+\epsilon)\right\}$, two transmissions again; or (iii) $\left\{\boldsymbol{\phi}_1^o(\epsilon), \boldsymbol{\phi}_1^1(\epsilon), \boldsymbol{\phi}_1^2(\epsilon)\right\}$, with three transmissions. Assume that (iii) attains the lowest distortion this causing the maximum number of transmissions (I=3) to be reached. From now on, no additional eigenvectors will be tried in subsequent iterations. However, some of the eigenvectors selected so far might need to be re-computed if any of the subsequent power tokens is allocated to a preceding one. The assumption here is that the greedy allocation of previous power tokens continues to be optimal for the re-computed eigenvectors, which is reasonable as long as ϵ is small.

The algorithm goes on until the L power tokens have been allocated. The (at most) I eigenvectors retained in the last iteration will be used as the actual set of pre-coders $\{\tilde{\mathbf{w}}_i\}_{i=1}^I$ along with the allocation of power tokens over such eigenvec-

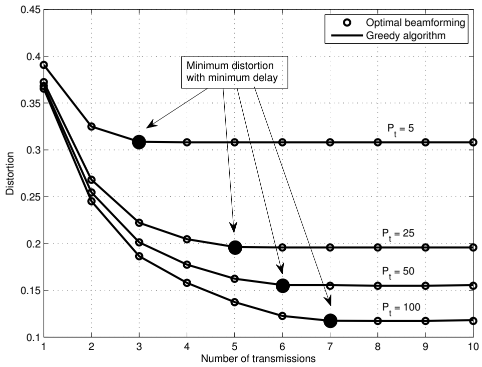

Fig. 3. Distortion vs transmission number (N=10, $\theta = 10^{-3}$)

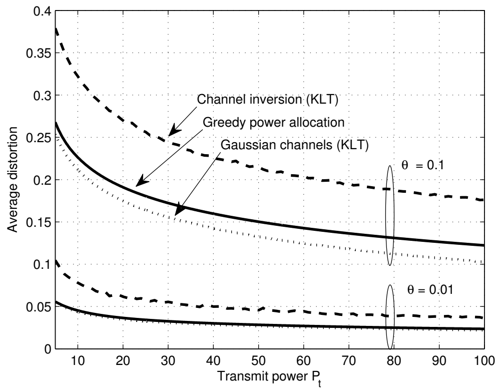

Fig. 4. Average distortion vs transmit power ($N = 25$)

tor set. Yet no optimality can be claimed for this approach, it exhibits a remarkable performance (see next section).

6. SIMULATION RESULTS AND CONCLUSIONS

The simulation scenario consists of N sensors deployed over a 10×10 rectangular area. As in [7], the spatial field is modeled as a Gaussian Markov Ornstein-Uhlenbeck process with correlation (covariance) function given by $k(\mathbf{s}_i, \mathbf{s}_j) = \sigma_x^2 \exp(-\theta \|\mathbf{s}_i - \mathbf{s}_j\|_2)$. In all cases, the variance of the spatial field and the additive noise read $\sigma_x^2 = 1$ and $\sigma_n^2 = 1$, respectively. Unless otherwise stated, the sensor-to-GW channels are assumed to be Rayleigh-fading.

Figure 3 shows some results for a setting with a random sensor deployment ($N = 10$ sensors, uniform distribution). First, we observe that the performance of the iterative (greedy) solution is virtually identical to the optimal one (numerically computed with Matlab). As it follows from (10), distortion decreases with the number of transmissions although beyond some point (big round markers on the curves) the curves saturate. This corresponds to the solution with the highest compression level (or, equivalently, lowest latency) for a given transmit power. In a practical implementation, no additional values of I would be searched for as soon as the decrease in distortion with respect to the previous value would be within a prescribed margin. By increasing the total transmit power available, a larger number of "useful" transmissions (pre-coders) can be afforded which effectively convey non-redundant information to the GW.

In Figure 4, we depict the reconstruction distortion averaged over channel realizations. Sensors here are deployed deterministically in a rectangular grid ($N = 25$ sensors in total). Unsurprisingly, distortion is lower when the field is highly correlated ($\theta = 0.01$). In this case, the available transmit power is allocated to a reduced number of pre-coders (since compression level can be higher) this resulting into a higher SNR per transmission. As a first benchmark we have used the scheme in our previous work [**?**] where a per-sensor channel equalization is carried out prior to applying a (distributed) Karhunen-Loeve transform to the set of observations (i.e. directly using the eigenvectors of the covariance matrix of the spatial field, $\mathbf{C}_x$). The use of the proposed successive refinement technique, by which knowledge on the statistical properties of the spatial field and the channel gains are *jointly* exploited for pre-coder design (rather than separately as in [**?**]) definitely pays off. As a second benchmark, we also depict the distortion attained in a scenario with Gaussian channels and the optimal pre-coding solution computed in Section 4.2. As expected, fading has a negative effect in terms of distortion since it has to be (partly) compensated for by the power allocation strategy.

In conclusion, the proposed iterative greedy scheme allows us to simultaneously (and effectively) solve the precoder design an power allocation problems for the general case. Performance is virtually identical to that of the optimal solution computed numerically. The gain with respect to other precoding schemes requiring per-sensor channel equalization is large.

7. REFERENCES

[1] "EXALTED." [Online]. Available: http://www.ict-exalted.eu/

[2] M. Gastpar, "Uncoded transmission is exactly optimal for a simple gaussian sensor network," *IEEE Trans. on Inf. Theory*, vol. 54, no. 11, pp. 5247 –5251, nov. 2008.

[3] L. Dong, A. Petropulu, and H. Poor, "Weighted cross-layer cooperative beamforming for wireless networks," *IEEE Transactions on Signal Processing*, vol. 57, no. 8, pp. 3240 –3252, aug. 2009.

[4] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Joint source channel communication for distributed estimation in sensor networks," *IEEE Trans. on Inf. Theory*, vol. 53, no. 10, pp. 3629 –3653, oct. 2007.

[5] S. Cui, J. Xiao, A. Goldsmith, Z.-Q. Luo, and H. V. Poor, "Estimation diversity and energy efficiency in distributed sensing," *IEEE Trans. on Signal Proc.*, vol. 55, no. 9, pp. 4683 – 4695, Sept. 2007.

[6] R. Mudumbai, J. Hespanha, U. Madhow, and G. Barriac, "Distributed transmit beamforming using feedback control," *IEEE Trans. on Inf. Theory*, vol. 56, no. 1, pp. 411 –426, jan. 2010.

[7] M. Dong, L. Tong, and B. Sadler, "Impact of data retrieval pattern on homogeneous signal field reconstruction in dense sensor networks," *IEEE Trans. on Signal Process.*, vol. 54, no. 11, pp. 4352–4364, Nov. 2006.