Decentralized Random-field Estimation for Sensor Networks Using Quantized Spatially Correlated Data and Fusion-center Feedback[†]

Aleksandar Dogandžić and *Kun Qiu*

ECpE Department, Iowa State University, 3119 Coover Hall, Ames, IA 50011

Phone: (515) 294-0500, Fax: (515) 294-8432

email: {ald,kqiu}@iastate.edu

Abstract

In large-scale wireless sensor networks, sensor-processor elements (nodes) are densely deployed to monitor the environment; consequently, their observations form a *random field* that is highly correlated in space. We consider a fusion sensor-network architecture where, due to the bandwidth and energy constraints, the nodes transmit quantized data to a fusion center. The fusion center provides feedback by broadcasting summary information to the nodes. In addition to saving energy, this feedback ensures reliability and robustness to node and fusion-center failures. We assume that the sensor observations follow a linear-regression model with known spatial covariances between any two locations within a region of interest. We propose a Bayesian framework for adaptive quantization, fusion-center feedback, and estimation of the random field and its parameters. We also derive a simple suboptimal scheme for estimating the unknown parameters, apply our estimation approach to the no-feedback scenario, discuss field prediction at arbitrary locations within the region of interest, and present numerical examples demonstrating the performance of the proposed methods.

I. INTRODUCTION

Decentralized parameter estimation for sensor networks has recently attracted considerable attention [1]–[5]. Due to the high density of deployed sensor-processor elements (nodes), the physical phenomena monitored by sensor networks (e.g. moisture content in an agricultural field, temperature distribution in a building, pH values quantifying ecological content of a river, or concentration of a chemical) yield observations that are highly correlated in space [6]–[11]. This spatial correlation has been utilized for estimation, detection, classification, quantization, data aggregation, compression, routing, querying, localization, and medium access control (MAC) protocol design, see [6]–[15] and references therein. In [4] and [5], efficient quantization strategies have been designed for estimating a constant-mean signal in spatially correlated Gaussian noise assuming that no feedback is available from a fusion center; in addition, [5] computes the corresponding Cramér-Rao bounds for the unknown parameters. Data-rate limited dynamic state estimation and control problems have also been studied in [16]–[18]. In [18], a Kalman-filter like distributed state estimation scheme is developed for the scenario where nodes transmit signs of innovations to the fusion

center. Kalman filtering in [19] considers missing observations due to unreliable network links. A framework for distributed estimation of deterministic spatial phenomena is proposed in [20], see also [21]. Utilizing and designing feedback links to improve communication and control over noisy channels have been studied in [22]–[24], see also references therein. Feedback from the fusion center to the nodes has been utilized to estimate states of finite-state Markov chains evolving in time [25] and space [15]. Decision feedback is studied in [26]–[29] in the context of hypothesis testing (detection). Autoregressive target tracking with feedback is proposed in [30], where the nodes transmit analog information to the fusion center. Estimation from faded misaligned observations using quantized feedback is studied in [10]. A decentralized quantile estimation scheme that utilizes quantized fusion-center feedback is proposed in [31]. However, fusion-center feedback has not yet been used for *adaptive quantization* at the nodes.

In this paper, we propose a Bayesian framework for adaptive quantization, fusion-center feedback, and estimation of a *spatial random field* and its parameters.

In Section II, we first introduce the measurement and prior models and then describe the proposed quantization and feedback schemes (Section II-A); a no-feedback scenario is outlined in Section II-B. A Monte-Carlo (MC) approach is developed in Section III for drawing independent, identically distributed (i.i.d.) samples from the posterior distribution of the random field; these samples are then utilized to estimate and predict the random field and its parameters. A simple suboptimal estimation approach and corresponding quantization and feedback schemes are proposed in Section IV. In Section V, we evaluate the performance of the proposed methods via numerical simulations. Concluding remarks are given in Section VI.

II. MEASUREMENT MODEL, PRIOR SPECIFICATION, AND QUANTIZATION AND FEEDBACK SCHEMES

Assume that a region of interest contains a cluster of N nodes at known locations $\boldsymbol{x}_n$, $n = 1, 2, \ldots, N$, measuring

$$y_n = \boldsymbol{h}_n^T \boldsymbol{\alpha} + e_n, \quad n = 1, 2, \ldots, N \tag{2.1a}$$

and a fusion center at location $\boldsymbol{x}_{N+1}$, measuring

$$y_{N+1} = \boldsymbol{h}_{N+1}^T \boldsymbol{\alpha} + e_{N+1} \tag{2.1b}$$

where "^{T}" denotes a transpose, $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_N, \boldsymbol{h}_{N+1}$ are known *regression vectors* of size $r \times 1$, and $\boldsymbol{\alpha}$ is an $r \times 1$ vector of unknown *regression coefficients*. Here,

- the first terms in (2.1a) and (2.1b) model *large-scale spatial-signal variation* (e.g. spatial trend) and

- the terms e_n, $n = 1, 2, \ldots, N + 1$ are zero-mean spatially correlated Gaussian random variables, modeling the *small-scale random-field variations*

see also [32, Ch. 9] and [33, Ch. 5.1]. This regression model and its special cases have been studied in [3]–[5] and [34]; it is fairly general and can incorporate signal representations based on splines, wavelets, Fourier series etc. For example, in Section V-B, we select the regression vectors as $\boldsymbol{h}_n = [1 \ \boldsymbol{x}_n^T]^T$, yielding a *linear trend-surface model* for the large-scale spatial-signal variation; see [5, Sect. VI.B] and [34] for other choices of $\boldsymbol{h}_n$ that have been employed in the sensor-network context. Define

$$\boldsymbol{v} = [\boldsymbol{y}_{1:N}^T, y_{N+1}]^T \tag{2.2a}$$

where

$$\boldsymbol{y}_{1:N} = [y_1, y_2, \ldots, y_N]^T \tag{2.2b}$$

is the vector of all measurements at the nodes within the region of interest. Denote by $\mathcal{N}(\boldsymbol{v}\,;\,\boldsymbol{\mu},\,\Psi)$ the multivariate probability density function (pdf) of a Gaussian random vector $\boldsymbol{v}$ with mean vector $\boldsymbol{\mu}$ and covariance matrix Ψ. Then, the pdf of $\boldsymbol{v}$ is multivariate Gaussian:

$$p(\boldsymbol{v}\,|\,\boldsymbol{\alpha}) = \mathcal{N}(\boldsymbol{v}\,;\,H\,\boldsymbol{\alpha},\,\Psi) \tag{2.2c}$$

where $\Psi = \mathrm{cov}(\boldsymbol{v}\,|\,\boldsymbol{\alpha})$ is the $(N{+}1)\times(N{+}1)$ *positive definite spatial covariance matrix* of the small-scale random-field component and

$$H = [\boldsymbol{h}_1\,\boldsymbol{h}_2\,\ldots\boldsymbol{h}_N\,\boldsymbol{h}_{N+1}]^T \tag{2.2d}$$

is the known regression matrix of the large-scale spatial-signal variation component. The variance of the small-scale random-field component is the (n, n) element of Ψ, denoted by

$$\psi_n^2 = \Psi_{n,n} = \mathrm{var}(y_n\,|\,\boldsymbol{\alpha}) = \mathrm{var}(e_n) \tag{2.2e}$$

for $n = 1, 2, \ldots, N + 1$.

Prior specification. We assign a (conjugate) Gaussian prior pdf for the regression parameters:

$$\pi(\boldsymbol{\alpha}) = \mathcal{N}(\boldsymbol{\alpha}\,;\,\boldsymbol{\mu_\alpha}, \Gamma_{\boldsymbol{\alpha}}) \tag{2.3}$$

where the $r \times 1$ vector $\boldsymbol{\mu_\alpha}$ and $r \times r$ covariance matrix $\Gamma_{\boldsymbol{\alpha}}$ quantify our prior knowledge about the large-scale spatial-signal variation. Here, the fusion center may employ its past estimates of $\boldsymbol{\alpha}$ to estimate $\boldsymbol{\mu_\alpha}$ and $\Gamma_{\boldsymbol{\alpha}}$.

A. Quantization and Feedback Schemes

We adopt a *fusion architecture* where the nodes $n = 1, 2, \ldots, N$ *coarsely quantize* their observations and convey the quantization results to the fusion center. This quantization conserves energy at the nodes, as well as radio-frequency (RF) bandwidth needed for communication from the nodes to the fusion center. We assume that the fusion center knows all regression vectors $\boldsymbol{h}_1, \boldsymbol{h}_2, \ldots, \boldsymbol{h}_N, \boldsymbol{h}_{N+1}$, the spatial covariance matrix Ψ of the small-scale random-field variations (which is also the assumption made in e.g. [4], [5], and [13]) and the prior mean $\boldsymbol{\mu}_{\boldsymbol{\alpha}}$ and covariance matrix $\Gamma_{\boldsymbol{\alpha}}$ for the regression-coefficient vector $\boldsymbol{\alpha}$. In contrast, each node $n = 1, 2, \ldots, N$ knows *only* its regression vector $\boldsymbol{h}_n$ and field variance ψ_n^2. (In many models, knowing $\boldsymbol{h}_n$ effectively reduces to knowing node n's location $\boldsymbol{x}_n$; hence, in such models, we require node n to know its location.)

The fact that the nodes utilize only their local information is an important property of our approach. Furthermore, the nodes do not have special arrangements with (i.e. are treated equally by) the fusion center and can continue to operate in the same manner even if the fusion center changes, e.g. switches from one node to another.

The fusion center utilizes the quantized data collected from the nodes and knowledge of the small-scale random-field covariances to estimate the regression parameters $\boldsymbol{\alpha}$, reconstruct the observations $\boldsymbol{y}_{1:N}$, and predict the random field at arbitrary locations within the region of interest. It also occasionally *broadcasts* summary information to the nodes, which we refer to as *fusion-center feedback*. This feedback facilitates energy-efficient node transmissions, see the numerical examples in Sections V-A.1 and V-B.1. It also ensures reliability and robustness to node and fusion-center failures, since all nodes within the region of interest receive summary information regarding the large-scale component of the observed spatial phenomenon.

To prepare for the first feedback broadcast, the fusion center determines the posterior pdf of $\boldsymbol{\alpha}$ based on its measurement y_{N+1} [see (2.1b) and (2.3)]:

$$p(\boldsymbol{\alpha} \mid y_{N+1}) = \mathcal{N}(\boldsymbol{\alpha}\,;\,\boldsymbol{\alpha}_{(0)}, C_{(0)}) \propto p(y_{N+1} \mid \boldsymbol{\alpha})\,\pi(\boldsymbol{\alpha}) \tag{2.4a}$$

where the mean and covariance of this Gaussian pdf are

$$\boldsymbol{\alpha}_{(0)} = C_{(0)} \left(\Gamma_{\boldsymbol{\alpha}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\alpha}} + \frac{\boldsymbol{h}_{N+1}\, y_{N+1}}{\psi_{N+1}^2} \right) \tag{2.4b}$$

$$C_{(0)} = \left(\Gamma_{\boldsymbol{\alpha}}^{-1} + \frac{\boldsymbol{h}_{N+1}\, \boldsymbol{h}_{N+1}^T}{\psi_{N+1}^2} \right)^{-1} = \Gamma_{\boldsymbol{\alpha}} - \frac{\Gamma_{\boldsymbol{\alpha}}\, \boldsymbol{h}_{N+1}\, \boldsymbol{h}_{N+1}^T\, \Gamma_{\boldsymbol{\alpha}}}{\psi_{N+1}^2 + \boldsymbol{h}_{N+1}^T\, \Gamma_{\boldsymbol{\alpha}}\, \boldsymbol{h}_{N+1}}. \tag{2.4c}$$

Here, $\boldsymbol{\alpha}_{(0)}$ is also the Bayesian minimum mean-square error (MMSE) estimate of $\boldsymbol{\alpha}$, given y_{N+1}.

We first describe the initial information exchange (referred to as *Round 1*) between the fusion center and the nodes $n = 1, 2, \ldots, N$ and then present an extension to multiple rounds.

Round 1. The fusion center broadcasts $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$ to the nodes within the region of interest.[1] We assume that sufficient RF bandwidth is available for these broadcasts and that the nodes within the region of interest receive them error-free. Upon receiving $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$, these nodes construct their local Gaussian prior pdf for $\boldsymbol{\alpha}$:

$$p(\boldsymbol{\alpha} \,|\, \boldsymbol{\alpha}_{(0)}, C_{(0)}) = \mathcal{N}(\boldsymbol{\alpha}\,;\, \boldsymbol{\alpha}_{(0)}, C_{(0)}) \tag{2.5a}$$

which, in the sequential-Bayesian spirit, *coincides* with the posterior pdf $p(\boldsymbol{\alpha} \,|\, y_{N+1})$ in (2.4a). Each node n utilizes the local prior (2.5a) to determine the *predictive pdf* of y_n based on $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$:

$$p(y_n \,|\, \boldsymbol{\alpha}_{(0)}, C_{(0)}) = \mathcal{N}(y_n\,;\, \boldsymbol{h}_n^T \boldsymbol{\alpha}_{(0)}, \boldsymbol{h}_n^T C_{(0)} \boldsymbol{h}_n + \psi_n^2). \tag{2.5b}$$

Observe that the local prior pdf (2.5a) is the same at all nodes, but the predictive pdfs (2.5b) for y_n differ (in general) from node to node. Now, node n determines its quantization thresholds $\tau_{n,(1)}(0) < \tau_{n,(1)}(1) < \cdots < \tau_{n,(1)}(K)$ as follows:

$$\tau_{n,(1)}(l) = q_{n,(1)}\left(\frac{l}{K}\right), \quad l = 0, 1, \ldots, K \tag{2.6a}$$

where $q_{n,(1)}(l/K)$ is the l/K quantile of the predictive pdf (2.5b) and K denotes the number of quantization intervals.[2] It then reports to the fusion center the index $k_{n,(1)} \in \{0, 1, 2, \ldots, K-1\}$ corresponding to the quantization interval $\mathcal{R}_{n,(1)}$ containing y_n:

$$\mathcal{R}_{n,(1)} = \left[\tau_{n,(1)}(k_{n,(1)}),\ \tau_{n,(1)}(1 + k_{n,(1)})\right). \tag{2.6b}$$

Hence, $k_{n,(1)} = j$ means that the measurement y_n falls within the range

$$\tau_{n,(1)}(j) \leq y_n < \tau_{n,(1)}(1 + j) \tag{2.6c}$$

and, consequently, $\mathcal{R}_{n,(1)} = \left[\tau_{n,(1)}(j),\ \tau_{n,(1)}(1 + j)\right)$; here, the first subscript "$_n$" corresponds to the node index and the second parenthesized subscript "$_{(1)}$" denotes the round index. In Round 1, the fusion center receives

$$\boldsymbol{k}_{(1)} = [k_{1,(1)}, k_{2,(1)}, \ldots, k_{N,(1)}]^T \tag{2.6d}$$

---

[1] In general, any node in the region of interest may serve as a *beacon*, compute $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$ based on its measurement and then broadcast them to other nodes or, alternatively, send its measurement to the fusion center. However, such schemes violate our assumption that are all nodes within the region of interest are treated equally by the fusion center. Therefore, we focus on the scenario where the fusion center uses its own measurement in Round 1 to determine $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$.

[2] Here, the trivial thresholds $\tau_{n,(1)}(0) = q_{n,(1)}(0) = -\infty$ and $\tau_{n,(1)}(K) = q_{n,(1)}(1) = +\infty$ are defined for notational convenience.

from the nodes in the region of interest. To interpret $\boldsymbol{k}_{(1)}$, the fusion center computes the same quantization thresholds as the nodes. For $K = 2$ quantization intervals, analog-to-digital conversion greatly simplifies and is reduced to a signal-level comparison [5]. In this case, the single nontrivial threshold is $\tau_{n,(1)}(1) = \boldsymbol{h}_n^T \boldsymbol{\alpha}_{(0)}$ and, therefore, the fusion center *does not* need to transmit $C_{(0)}$ to the nodes.

The fusion center completes Round 1 by estimating the posterior mean vector $\boldsymbol{\alpha}_{(1)} = \mathrm{E}\left[\boldsymbol{\alpha} \mid y_{N+1}, \mathcal{R}_{(1)}\right]$ and covariance matrix $C_{(1)} = \mathrm{cov}(\boldsymbol{\alpha} \mid y_{N+1}, \mathcal{R}_{(1)})$ of the regression parameters, where $\mathrm{E}[\cdot]$ denotes expectation and $\mathcal{R}_{(1)} = (\mathcal{R}_{1,(1)}, \mathcal{R}_{2,(1)}, \ldots, \mathcal{R}_{N,(1)})$ is the collection of all quantization intervals obtained by the fusion center in Round 1; see also Section III. If Round 2 is planned, each node n saves its first-round quantization interval $\mathcal{R}_{n,(1)}$, to be used for *adaptive quantization*, as discussed in the following.

Round $p > 1$. Suppose that, after Round 1, the fusion center decides to continue its information exchange with the nodes $n = 1, 2, \ldots, N$. This exchange proceeds as follows:

- The fusion center *broadcasts* $\boldsymbol{\alpha}_{(p-1)}$ and $C_{(p-1)}$ (estimated at the end of Round $p - 1$) to the nodes $n = 1, 2, \ldots, N$. Upon receiving $\boldsymbol{\alpha}_{(p-1)}$ and $C_{(p-1)}$, these nodes *update* their local Gaussian prior pdf for $\boldsymbol{\alpha}$:

$$p(\boldsymbol{\alpha} \mid \boldsymbol{\alpha}_{(p-1)}, C_{(p-1)}) = \mathcal{N}(\boldsymbol{\alpha} ; \boldsymbol{\alpha}_{(p-1)}, C_{(p-1)}). \tag{2.7}$$

- To determine node n's thresholds in Round p, we should exploit available information that this node provided to the fusion center over the past $p - 1$ rounds. This information is quantified by an interval denoted by $\mathcal{R}_{n,(p-1)}$ [computed recursively using (2.9c)]. Note that (2.7) approximates the analytically intractable marginal posterior pdf $p(\boldsymbol{\alpha} \mid y_{N+1}, \mathcal{R}_{(p-1)})$. Node n utilizes (2.7) and $\mathcal{R}_{n,(p-1)}$ to determine the predictive pdf of y_n based on $\boldsymbol{\alpha}_{(p-1)}$, $C_{(p-1)}$, and $\mathcal{R}_{n,(p-1)}$:

$$p(y_n \mid \boldsymbol{\alpha}_{(p-1)}, C_{(p-1)}, \mathcal{R}_{n,(p-1)}) \propto \mathcal{N}(y_n ; \boldsymbol{h}_n^T \boldsymbol{\alpha}_{(p-1)}, \boldsymbol{h}_n^T C_{(p-1)} \boldsymbol{h}_n + \psi_n^2) \cdot i_{\mathcal{R}_{n,(p-1)}}(y_n) \tag{2.8a}$$

where

$$i_A(x) = \begin{cases} 1, & x \in A, \\ 0, & \text{otherwise} \end{cases} \tag{2.8b}$$

denotes the indicator function. Now, node n determines its quantization thresholds $\tau_{n,(p)}(0) < \tau_{n,(p)}(1) < \cdots < \tau_{n,(p)}(K)$ as follows:

$$\tau_{n,(p)}(l) = q_{n,(p)}\left(\frac{l}{K}\right), \quad l = 0, 1, \ldots, K \tag{2.9a}$$

where $q_{n,(p)}(l/K)$ is the l/K quantile of the predictive pdf (2.8a). It then reports to the fusion center the index $k_{n,(p)} \in \{0, 1, 2, \ldots, K - 1\}$ corresponding to the quantization interval $\left[\tau_{n,(p)}(k_{n,(p)}), \tau_{n,(p)}(1 + \right.$
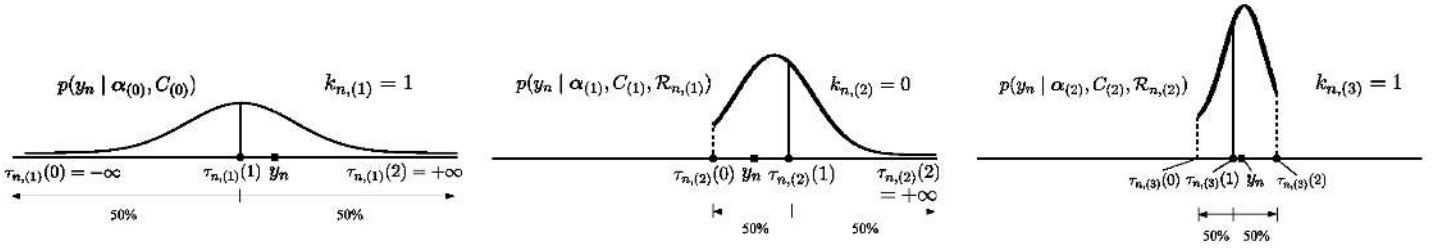
Fig. 1. Rounds 1,2, and 3 of the proposed adaptive quantization scheme, for $K = 2$ quantization intervals.

$k_{n,(p)}$)) containing y_n. Hence, in Round p, the fusion center receives

$$\boldsymbol{k}_{(p)} = [k_{1,(p)}, k_{2,(p)}, \ldots, k_{N,(p)}]^T. \tag{2.9b}$$

To interpret $\boldsymbol{k}_{(p)}$, the fusion center computes the same quantization thresholds as the nodes.

- Upon receiving $\boldsymbol{k}_{(p)}$, the fusion center determines $\mathcal{R}_{n,(p)}$ using the recursive formula:

$$\mathcal{R}_{n,(p)} = \mathcal{R}_{n,(p-1)} \cap \left[\tau_{n,(p)}(k_{n,(p)}), \tau_{n,(p)}(1 + k_{n,(p)})\right). \tag{2.9c}$$

If Round $p + 1$ is planned, each node n computes its $\mathcal{R}_{n,(p)}$ using (2.9c), to be used for adaptive quantization in Round $p + 1$. Now,

$$\mathcal{R}_{(p)} = (\mathcal{R}_{1,(p)}, \mathcal{R}_{2,(p)}, \ldots, \mathcal{R}_{N,(p)}) \tag{2.9d}$$

quantifies the fusion-center's knowledge of the intervals within which y_n, $n = 1, 2, \ldots, N$ fall based on the information provided by the nodes over the past p rounds.

- The fusion center completes Round p by estimating $\boldsymbol{\alpha}_{(p)}$ and $C_{(p)}$, as shown in Section III.

Our adaptive quantization scheme is depicted in Fig. 1.

B. No Feedback

If there is no fusion-center feedback, we determine the K quantization intervals using the following *fixed thresholds*:

$$\tau_{n,(1)}^{\text{NF}}(l) = q_{n,(1)}^{\text{NF}}\left(\frac{l}{K}\right), \quad l = 0, 1, \ldots, K \tag{2.10a}$$

where $q_{n,(1)}^{\text{NF}}(l/K)$ is the l/K quantile of the the *marginal pdf* of y_n:

$$p(y_n) = \mathcal{N}\left(y_n \, ; \, \boldsymbol{h}_n^T \boldsymbol{\mu_\alpha}, \boldsymbol{h}_n^T \Gamma_{\boldsymbol{\alpha}} \boldsymbol{h}_n + \psi_n^2\right). \tag{2.10b}$$

Note that (2.10b) can be obtained by replacing $\boldsymbol{\alpha}_{(0)}$ and $C_{(0)}$ in (2.5b) with the prior mean $\boldsymbol{\mu_\alpha}$ and covariance matrix $\Gamma_{\boldsymbol{\alpha}}$ for the regression-coefficient vector $\boldsymbol{\alpha}$. The nodes $n = 1, 2, \dots, N$ employ the thresholds (2.10a) to quantize their measurements and then transmit the quantized information to the fusion center, completing *Transmission 1*. Here, in addition to its regression vector $\boldsymbol{h}_n$ and field variance ψ_n^2, each node $n = 1, 2, \dots, N$ must know $\boldsymbol{\mu_\alpha}$ and $\Gamma_{\boldsymbol{\alpha}}$, which may be obtained during the initial network setup. We can continue and have multiple "no feedback" transmissions: in Transmission p, each node n quantizes its measurement y_n in a manner analogous to our quantization scheme from Section II-A, with $\boldsymbol{\alpha}_{(p)}$ and $C_{(p)}$ replaced by $\boldsymbol{\mu_\alpha}$ and $\Gamma_{\boldsymbol{\alpha}}$, respectively, for all $p \geq 0$. If we employ censoring in each transmission from the nodes, multiple transmissions will reduce average transmission energy compared with an equivalent (single) "no feedback" transmission. Censoring is also used for energy-efficient detection in sensor networks [35].

III. Estimation of $\boldsymbol{\alpha}_{(p)}$ and $\boldsymbol{C}_{(p)}$ and Random-field Estimation and Prediction

At the end of Round p, the fusion center selects $\boldsymbol{\alpha}_{(p)}$ and $C_{(p)}$ as the marginal posterior mean vector and covariance matrix of the regression parameters:

$$\boldsymbol{\alpha}_{(p)} = \mathrm{E}\left[\boldsymbol{\alpha} \,|\, y_{N+1}, \mathcal{R}_{(p)}\right], \quad C_{(p)} = \mathrm{cov}(\boldsymbol{\alpha} \,|\, y_{N+1}, \mathcal{R}_{(p)}). \tag{3.1}$$

Note that $\boldsymbol{\alpha}_{(p)}$ and $C_{(p)}$ cannot be determined in closed form [except when $p = 0$, see (2.4b)–(2.4c)]; hence, we estimate them via MC sampling. Let us first obtain the pdf $p(\boldsymbol{v}, \boldsymbol{\alpha} \,|\, \mathcal{R}_{(p)})$:

$$p(\boldsymbol{v}, \boldsymbol{\alpha} \,|\, \mathcal{R}_{(p)}) \propto p(\boldsymbol{v} \,|\, \boldsymbol{\alpha})\, \pi(\boldsymbol{\alpha}) \cdot \prod_{n=1}^{N} i_{\mathcal{R}_{n,(p)}}(y_n) \tag{3.2a}$$

see (2.2c), (2.3), and (2.9d). Then, the kernel of the conditional posterior pdf of $\boldsymbol{\alpha}$ given $\boldsymbol{v}$ and $\mathcal{R}_{(p)}$ is

$$\begin{aligned} p(\boldsymbol{\alpha} \,|\, \boldsymbol{v}, \mathcal{R}_{(p)}) &= p(\boldsymbol{\alpha} \,|\, \boldsymbol{v}) \propto p(\boldsymbol{v} \,|\, \boldsymbol{\alpha})\, \pi(\boldsymbol{\alpha}) \\ &\propto \exp[-\tfrac{1}{2}(\boldsymbol{v} - H\boldsymbol{\alpha})^T\, \Psi^{-1}(\boldsymbol{v} - H\boldsymbol{\alpha})] \cdot \exp[-\tfrac{1}{2}(\boldsymbol{\alpha} - \boldsymbol{\mu_\alpha})^T\, \Gamma_{\boldsymbol{\alpha}}^{-1}(\boldsymbol{\alpha} - \boldsymbol{\mu_\alpha})] \end{aligned} \tag{3.2b}$$

implying that $p(\boldsymbol{\alpha} \,|\, \boldsymbol{v}, \mathcal{R}_{(p)})$ is a multivariate Gaussian pdf:

$$p(\boldsymbol{\alpha} \,|\, \boldsymbol{v}, \mathcal{R}_{(p)}) = \mathcal{N}\left(\boldsymbol{\alpha};\, \Omega\,(\Gamma_{\boldsymbol{\alpha}}^{-1}\, \boldsymbol{\mu_\alpha} + H^T\, \Psi^{-1}\boldsymbol{v}),\, \Omega\right) \tag{3.2c}$$

where

$$\Omega = (\Gamma_{\boldsymbol{\alpha}}^{-1} + H^T\, \Psi^{-1}H)^{-1}. \tag{3.2d}$$

The marginal posterior pdf of $\boldsymbol{v}$ given $\mathcal{R}_{(p)}$ is a *truncated multivariate Gaussian*:

$$p(\boldsymbol{v} \,|\, \mathcal{R}_{(p)}) \propto p(\boldsymbol{v}) \cdot \prod_{n=1}^{N} i_{\mathcal{R}_{n,(p)}}(y_n) \tag{3.3a}$$

where $p(\boldsymbol{v})$ denotes the marginal pdf of $\boldsymbol{v}$:

$$p(\boldsymbol{v}) \; = \; \mathcal{N}(H\,\boldsymbol{\mu_\alpha}, \varSigma) \tag{3.3b}$$

$$\varSigma \; = \; \text{cov}(\boldsymbol{v}) = H\,\Gamma_{\boldsymbol{\alpha}}\,H^T + \Psi. \tag{3.3c}$$

The marginal variance of y_n is the (n,n) element of $\varSigma$, denoted by

$$\sigma_n^2 = \varSigma_{n,n} = \text{var}(y_n), \quad n = 1, 2, \ldots, N+1. \tag{3.3d}$$

We also adopt the following block partitioning of $\varSigma$:

$$\varSigma = \text{cov}(\boldsymbol{v}) = \begin{bmatrix} \varSigma_{1:N,1:N} & \boldsymbol{\sigma}_{1:N,N+1} \\ \boldsymbol{\sigma}_{1:N,N+1}^T & \sigma_{N+1}^2 \end{bmatrix} \tag{3.3e}$$

where $\varSigma_{1:N,1:N} = \text{cov}(\boldsymbol{y}_{1:N})$ and $\boldsymbol{\sigma}_{1:N,N+1} = \text{cov}(\boldsymbol{y}_{1:N}, y_{N+1})$ are $N \times N$ matrix and $N \times 1$ column vector, respectively.

Using the conditional pdf result for multivariate Gaussian distributions in e.g. [36, Theorem 10.2], we obtain an expression for the kernel of the marginal posterior pdf $p(\boldsymbol{y}_{1:N} \,|\, y_{N+1}, \mathcal{R}_{(p)})$:

$$p(\boldsymbol{y}_{1:N} \,|\, y_{N+1}, \mathcal{R}_{(p)}) \propto \mathcal{N}(\boldsymbol{y}_{1:N}\,;\,\boldsymbol{m}, S) \cdot \prod_{n=1}^{N} i_{\mathcal{R}_{n,(p)}}(y_n) \tag{3.4a}$$

where

$$\boldsymbol{m} \; = \; [\boldsymbol{h}_1\,\boldsymbol{h}_2 \cdots \boldsymbol{h}_N]^T\,\boldsymbol{\mu_\alpha} + \frac{\boldsymbol{\sigma}_{1:N,N+1}}{\sigma_{N+1}^2} \cdot (y_{N+1} - \boldsymbol{h}_{N+1}^T\,\boldsymbol{\mu_\alpha}) \tag{3.4b}$$

$$S \; = \; \varSigma_{1:N,1:N} - \frac{\boldsymbol{\sigma}_{1:N,N+1}\,\boldsymbol{\sigma}_{1:N,N+1}^T}{\sigma_{N+1}^2}. \tag{3.4c}$$

In the following section, we develop an algorithm for drawing i.i.d. samples from this distribution. The obtained samples are then utilized to calculate the marginal posterior mean and covariance matrix of $\boldsymbol{\alpha}$ in (3.1) and to estimate and predict the random field within the region of interest.

A. Random-field Sampler and Signal-parameter and Random-field Estimators

We now present a scheme for sampling $\boldsymbol{y}_{1:N,(p)}^{(t)}$, $t = 1, 2, \ldots, T$ from the marginal posterior pdf in (3.4a). First, decompose the $N \times N$ positive definite covariance matrix S in (3.4c) as follows:

$$S = \mathcal{L}\,\mathcal{L}^T \tag{3.5a}$$

where $\mathcal{L}$ is the lower-triangular Cholesky root of S:

$$\mathcal{L} = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \ldots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ l_{N1} & l_{N2} & l_{N3} & \cdots & l_{NN} \end{bmatrix}. \tag{3.5b}$$

Denote the nth element of $\mathcal{L}^{-1}\boldsymbol{m}$ by $[\mathcal{L}^{-1}\boldsymbol{m}]_n$. We obtain $\boldsymbol{y}_{1:N,(p)}^{(t)}$ as follows:

1. First, draw $x_{1,(p)}^{(t)}$ from

$$p(x_1\,|\,y_{N+1}, \mathcal{R}_{(p)}) \propto \mathcal{N}(x_1\,;\,[\mathcal{L}^{-1}\boldsymbol{m}]_1, 1)\cdot i_{\mathcal{R}_{1,(p)}}(l_{11}\,x_1); \tag{3.6a}$$

2. draw $x_{2,(p)}^{(t)}$ from

$$p(x_2\,|\,x_{1,(p)}^{(t)}, y_{N+1}, \mathcal{R}_{(p)}) \;\propto\; \mathcal{N}(x_2\,;\,[\mathcal{L}^{-1}\boldsymbol{m}]_2, 1)\cdot i_{\mathcal{R}_{2,(p)}}(l_{21}\,x_{1,(p)}^{(t)} + l_{22}\,x_2); \tag{3.6b}$$

$\cdots$

N. draw $x_{N,(p)}^{(t)}$ from

$$p(x_N\,|\,x_{1,(p)}^{(t)}, x_{2,(p)}^{(t)}, \ldots, x_{N-1,(p)}^{(t)}, y_{N+1}, \mathcal{R}_{(p)})$$
$$\propto \mathcal{N}(x_N\,;\,[\mathcal{L}^{-1}\boldsymbol{m}]_N, 1)\cdot i_{\mathcal{R}_{N,(p)}}(l_{N1}\,x_{1,(p)}^{(t)} + \cdots + l_{N(N-1)}\,x_{N-1,(p)}^{(t)} + l_{NN}\,x_N). \tag{3.6c}$$

Now, compute

$$\boldsymbol{y}_{1:N,(p)}^{(t)} = \mathcal{L}\,[x_{1,(p)}^{(t)}, x_{2,(p)}^{(t)}, \ldots, x_{N-1,(p)}^{(t)}, x_{N,(p)}^{(t)}]^T \tag{3.6d}$$

and define

$$\boldsymbol{v}_{(p)}^{(t)} = [(\boldsymbol{y}_{1:N,(p)}^{(t)})^T, y_{N+1}]^T. \tag{3.7}$$

The steps $1, 2, \ldots, N$ require sampling from truncated univariate Gaussian pdfs, which we describe in Appendix A. The sampling scheme (3.6) is more efficient than the Gibbs sampler in [37, Sect. 3] and [38, Sect. 3], where sampling from full conditional pdfs of y_n was used to perform an equivalent task.

Random-field estimation in Round p. Once we have collected the desired number of samples T, we estimate $\boldsymbol{y}_{1:N}$ by averaging the random-field draws:

$$\overline{\boldsymbol{y}}_{1:N,(p)} = \mathrm{E}\,[\boldsymbol{y}_{1:N}\,|\,y_{N+1}, \mathcal{R}_{(p)}] \approx \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{y}_{1:N,(p)}^{(t)}. \tag{3.8a}$$

Note that (3.9a) and (3.8a) can be viewed as (approximate) Round p's MMSE estimates of $\boldsymbol{\alpha}$ and $\boldsymbol{y}_{1:N}$. Define

$$\overline{\boldsymbol{v}}_{(p)} = [\overline{\boldsymbol{y}}_{1:N,(p)}^T, y_{N+1}]^T. \tag{3.8b}$$

Regression-parameter estimation in Round p. To estimate the marginal posterior mean vector and covariance matrix of $\boldsymbol{\alpha}$ in (3.1) when $p \geq 1$, we employ Rao-Blackwellization[3] based on (3.2c):

$$
\begin{aligned}
\boldsymbol{\alpha}_{(p)} &\approx \frac{1}{T} \sum_{t=1}^{T} (\Omega\, \Gamma_{\boldsymbol{\alpha}}^{-1}\, \boldsymbol{\mu}_{\boldsymbol{\alpha}} + \boldsymbol{\omega}_{(p)}^{(t)}) \\
&= \Omega\, \Gamma_{\boldsymbol{\alpha}}^{-1}\, \boldsymbol{\mu}_{\boldsymbol{\alpha}} + \overline{\boldsymbol{\omega}}_{(p)}
\end{aligned} \tag{3.9a}
$$

$$
C_{(p)} \approx \Omega + \left[\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{\omega}_{(p)}^{(t)} (\boldsymbol{\omega}_{(p)}^{(t)})^{T} \right] - \overline{\boldsymbol{\omega}}_{(p)}\, \overline{\boldsymbol{\omega}}_{(p)}^{T} \tag{3.9b}
$$

where Ω was defined in (3.2d) and

$$
\boldsymbol{\omega}_{(p)}^{(t)} = \Omega\, H^{T}\, \Psi^{-1}\, \boldsymbol{v}_{(p)}^{(t)} \tag{3.9c}
$$

$$
\overline{\boldsymbol{\omega}}_{(p)} = \Omega\, H^{T}\, \Psi^{-1}\, \overline{\boldsymbol{v}}_{(p)}. \tag{3.9d}
$$

In particular, Rao-Blackwellization based on (3.2c) yields the following estimate of the marginal posterior pdf of $\boldsymbol{\alpha}$:

$$
p(\boldsymbol{\alpha}|y_{N+1}, \mathcal{R}_{(p)}) \approx \frac{1}{T} \sum_{t=1}^{T} \mathcal{N}\big(\boldsymbol{\alpha};\, \Omega \Gamma_{\boldsymbol{\alpha}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\alpha}} + \boldsymbol{\omega}_{(p)}^{(t)},\, \Omega\big) \tag{3.9e}
$$

and (3.9a)–(3.9b) follow from this pdf. Clearly, $C_{(p)} - \Omega$ is a positive semidefinite matrix, implying that $C_{(p)}$ is lower-bounded by Ω. A single sample $T = 1$ in (3.9b) leads to the crude covariance estimate $C_{(p)} \approx \Omega$.

B. Random-field Prediction

Our sampler can be utilized to predict the random field y_i at a location $\boldsymbol{x}_i$ where sensors *have not* been deployed. Here, we are motivated by the fact that a typical goal of a sensor network is to infer the state of nature from *partial* and *noisy* information about the phenomenon of interest, which is in contrast with simple recovery of sensor readings by the fusion center. To predict the random field well, we must exploit the correlation structure of the small-scale random-field component. We assume that the random-field value y_i follows the model (2.1a):

$$
y_i = \boldsymbol{h}_i^{T}\, \boldsymbol{\alpha} + e_i \tag{3.10a}
$$

and that the fusion center *knows* the corresponding regression vector $\boldsymbol{h}_i$ and small-scale random-field variance and covariances

$$
\psi_i^{2} = \operatorname{var}(y_i \,|\, \boldsymbol{\alpha}) = \operatorname{var}(e_i) \tag{3.10b}
$$

$$
\boldsymbol{\psi}_i = \operatorname{cov}(\boldsymbol{v}, y_i \,|\, \boldsymbol{\alpha}). \tag{3.10c}
$$

---

[3]Rao-Blackwellization is a well-known variance-reduction method applied in Monte-Carlo computations, see [39, Chs. 2.3 and 6.6.3] and references therein.

Using the conditional pdf result for multivariate Gaussian distributions in [36, Theorem 10.2], we obtain an expression for the conditional posterior pdf $p(y_i \,|\, \boldsymbol{v}, \mathcal{R}_{(p)}, \boldsymbol{\alpha})$:

$$
\begin{aligned}
p(y_i \,|\, \boldsymbol{v}, \mathcal{R}_{(p)}, \boldsymbol{\alpha}) &= p(y_i \,|\, \boldsymbol{v}, \boldsymbol{\alpha}) \\
&= \mathcal{N}\big(y_i \,;\, \boldsymbol{q}_i^T \boldsymbol{\alpha} + \boldsymbol{\psi}_i^T \, \Psi^{-1} \, \boldsymbol{v}, \psi_i^2 - \boldsymbol{\psi}_i^T \, \Psi^{-1} \boldsymbol{\psi}_i\big)
\end{aligned}
\tag{3.11a}
$$

where

$$
\boldsymbol{q}_i = \boldsymbol{h}_i - H^T \, \Psi^{-1} \, \boldsymbol{\psi}_i.
\tag{3.11b}
$$

Using (3.2c) and (3.11a), we obtain the marginal posterior-predictive pdf

$$
p(y_i \,|\, \boldsymbol{v}, \mathcal{R}_{(p)}) = p(y_i \,|\, \boldsymbol{v}) = \mathcal{N}\Big(y_i \,;\, \boldsymbol{q}_i^T \, \Omega \, (\Gamma_{\boldsymbol{\alpha}}^{-1} \boldsymbol{\mu}_{\boldsymbol{\alpha}} + H^T \, \Psi^{-1} \, \boldsymbol{v}) + \boldsymbol{\psi}_i^T \, \Psi^{-1} \boldsymbol{v}, \mathrm{var}(y_i | \boldsymbol{v})\Big)
\tag{3.12a}
$$

where

$$
\mathrm{var}(y_i \,|\, \boldsymbol{v}) = \boldsymbol{q}_i^T \, \Omega \, \boldsymbol{q}_i + \psi_i^2 - \boldsymbol{\psi}_i^T \, \Psi^{-1} \boldsymbol{\psi}_i.
\tag{3.12b}
$$

Now, Rao-Blackwellization based on (3.12a) yields the following predictor of y_i and corresponding posterior-predictive variance:

$$
\overline{y}_{i,(p)} = \mathrm{E}\left[y_i | y_{N+1}, \mathcal{R}_{(p)}\right] \approx \boldsymbol{q}_i^T \boldsymbol{\alpha}_{(p)} + \boldsymbol{\psi}_i^T \, \Psi^{-1} \overline{\boldsymbol{v}}_{(p)}
\tag{3.13a}
$$

$$
\mathrm{var}(y_i \,|\, y_{N+1}, \mathcal{R}_{(p)}) \approx \mathrm{var}(y_i \,|\, \boldsymbol{v}) + (H \, \Omega \, \boldsymbol{q}_i + \boldsymbol{\psi}_i)^T \, \Psi^{-1} \, \Xi \, \Psi^{-1} \, (H \, \Omega \, \boldsymbol{q}_i + \boldsymbol{\psi}_i)
\tag{3.13b}
$$

where

$$
\Xi = \left[\frac{1}{T} \sum_{t=1}^{T} \boldsymbol{v}_{(p)}^{(t)} \, (\boldsymbol{v}_{(p)}^{(t)})^T\right] - \overline{\boldsymbol{v}}_{(p)} \, \overline{\boldsymbol{v}}_{(p)}^T
\tag{3.13c}
$$

and $\boldsymbol{v}_{(p)}^{(t)}$, $\overline{\boldsymbol{v}}_{(p)}$, and $\boldsymbol{\alpha}_{(p)}$ are computed using (3.7), (3.8b), and (3.9a).

Random-field estimation as a special case. For $n = 1, 2, \ldots, N$, $\boldsymbol{q}_n = \boldsymbol{0}_{r \times 1}$, where $\boldsymbol{0}_{r \times 1}$ denotes the $r \times 1$ vector of zeros. Consequently, (3.13a) with $i = n$ reduces to the nth element of $\overline{\boldsymbol{y}}_{1:N,(p)}$ in (3.8a) and (3.13b) reduces to the (n, n) element of Ξ:

$$
\mathrm{var}(y_n \,|\, y_{N+1}, \mathcal{R}_{(p)}) = \Xi_{n,n}.
\tag{3.14}
$$

We now outline a simple suboptimal estimation approach and corresponding quantization and feedback schemes.

IV. A SUBOPTIMAL SCHEME FOR ESTIMATION, QUANTIZATION, AND FEEDBACK

We propose a suboptimal method for estimating the field measurements y_n and regression coefficients $\boldsymbol{\alpha}$ in Round p. We mimic the sampling scheme in Section III-A and generate a single *deterministic sample* of the random field as follows:

$$\boldsymbol{y}_{1:N,(p)}^{(0)} = \mathcal{L}\,[x_{1,(p)}^{(0)}, x_{2,(p)}^{(0)}, \ldots, x_{N-1,(p)}^{(0)}, x_{N,(p)}^{(0)}]^T \tag{4.1a}$$

where we use the parenthesized superscript index "$^{(0)}$" to distinguish this deterministic sample from the random MC samples, which have indices $t = 1, 2, \ldots, T$. Here, $x_{n,(p)}^{(0)}$, $n = 1, 2, \ldots, N$ are recursively computed as

$$x_{n,(p)}^{(0)} = \mathrm{E}\,\big[x_N \,\big|\, x_{1,(p)}^{(0)}, x_{2,(p)}^{(0)}, \ldots, x_{N-1,(p)}^{(0)}, y_{N+1}, \mathcal{R}_{(p)}\big] \tag{4.1b}$$

i.e. $x_{n,(p)}^{(0)}$ are the means of the conditional pdfs that we sampled from in our MC scheme (3.6) (with t set to zero to indicate deterministic sampling), see also Appendix B. After computing $\boldsymbol{y}_{1:N,(p)}^{(0)}$, we estimate the regression parameters as

$$\boldsymbol{\alpha}_{(p)}^{(0)} = \Omega\,(\Gamma_{\boldsymbol{\alpha}}^{-1}\,\boldsymbol{\mu}_{\boldsymbol{\alpha}} + H^T\,\Psi^{-1}\boldsymbol{v}_{(p)}^{(0)}) \tag{4.1c}$$

where $\boldsymbol{v}_{(p)}^{(0)} = [(\boldsymbol{y}_{1:N,(p)}^{(0)})^T, y_{N+1}]^T$. This $\boldsymbol{\alpha}_{(p)}^{(0)}$ is the standard linear MMSE estimator, with $\boldsymbol{v}$ replaced by its estimate $\boldsymbol{v}_{(p)}^{(0)}$; it is also equal to the Rao-Blackwellized estimate in (3.9a) with $T = 1$ and t set to zero to indicate deterministic sampling.

Quantization. In Round 1, we select $\boldsymbol{\alpha}_{(0)}^{(0)}, C_{(0)}^{(0)}$, and $\mathcal{R}_{n,(1)}$ identical to (2.4b), (2.4c), and (2.6b), respectively. In Rounds $p > 1$, we choose $C_{(p-1)}^{(0)}$ as the following crude estimate of the posterior covariance matrix of $\boldsymbol{\alpha}$:

$$C_{(1)}^{(0)} = C_{(2)}^{(0)} = \ldots = \Omega \tag{4.1d}$$

equal to our Rao-Blackwellized estimate in (3.9b) with $T = 1$. Note that Ω needs to be transmitted to the nodes only once, during the initial network setup. The adaptive quantization proceeds in the same manner as described in Section II-A, with $\boldsymbol{\alpha}_{(p)}$ and $C_{(p)}$ replaced by $\boldsymbol{\alpha}_{(p)}^{(0)}$ and $C_{(p)}^{(0)}$.

Suboptimal random-field prediction in Round p. To predict y_i, we apply (3.13a) with $\boldsymbol{\alpha}_{(p)}$ and $\overline{\boldsymbol{v}}_{(p)}$ replaced by $\boldsymbol{\alpha}_{(p)}^{(0)}$ and $[(\boldsymbol{y}_{1:N,(p)}^{(0)})^T, y_{N+1}]^T$, respectively. Denote the resulting suboptimal predictor by $y_{i,(p)}^{(0)}$.

V. Numerical Examples

To assess the performance of the proposed methods, we consider a sensor network with N nodes randomly (uniformly) placed on a 10 m $\times$ 10 m two-dimensional (2-D) grid with 1 m spacing between the potential node locations. The fusion center was placed at the center of the grid. We simulated the observations $\boldsymbol{v}$ using (2.2c) where we selected an *isotropic exponential spatial covariance model* for the small-scale random-field component, with (see e.g. [6], [13], and [33, Ch. 2.1.3]):

$$\mathrm{cov}(y_i, y_j \,|\, \boldsymbol{\alpha}) = \mathrm{cov}(e_i, e_j) = \psi^2 \cdot \exp(-\phi \, \|\boldsymbol{x}_i - \boldsymbol{x}_j\|). \tag{5.1}$$

Here, the locations $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are 2-D Cartesian-coordinate column vectors, $\|\cdot\|$ denotes the Euclidean norm (i.e. $\|\boldsymbol{x}\|^2 = \boldsymbol{x}^T \boldsymbol{x}$), and $\psi^2 = \mathrm{var}(y_n \,|\, \boldsymbol{\alpha}) = \mathrm{var}(e_n)$ and $\phi > 0$ are known constants. Note that ψ^2 is the variance of the small-scale random-field component and ϕ quantifies the spatial-correlation strength of the field, with large values of ϕ corresponding to low spatial correlation and vice versa. We set the number of nodes to

$$N = 10 \tag{5.2a}$$

and the small-scale random-field variance to

$$\psi^2 = 1 \tag{5.2b}$$

except in Section V-B.2, where we vary ψ^2 (Figs. 11 and 12).

A. Constant Mean Signal

We adopt the constant-mean signal model with $\boldsymbol{h}_n = 1$:

$$H = \boldsymbol{1}_{N+1} \tag{5.3a}$$

where $\boldsymbol{1}_{N+1}$ denotes the $(N+1) \times 1$ vector of ones. In this case, $\boldsymbol{\alpha} = \alpha$ is the *constant mean signal*. Here, we choose a diffuse prior pdf for α, with (see (2.3))

$$\mu_\alpha = 0, \quad \Gamma_\alpha = 25 \tag{5.3b}$$

implying that the marginal variance of the measurements y_n is

$$\mathrm{var}(y_n) = \Sigma_{n,n} = \Gamma_\alpha + \psi^2 = 26 \tag{5.3c}$$

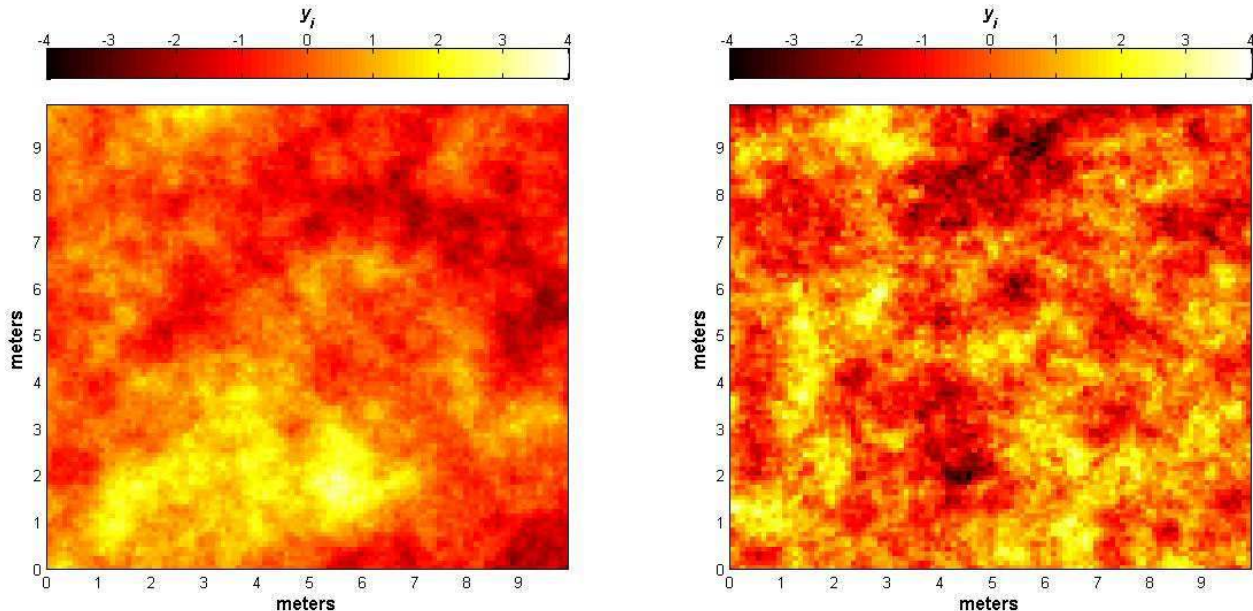which quantifies the dynamic range of measurements at the nodes.

Fig. 2. Constant mean-signal scenario: Random-field realizations for (a) $\phi = 0.5$ m^{-1} and (b) $\phi = 2$ m^{-1}.

In Fig. 2, we show random-field realizations generated for the correlation model (5.1) and mean signal $\alpha = 0$ using the method in [40]. To demonstrate the importance of the correlation-strength parameter ϕ, we generated two random-field realizations, with (a) $\phi = 0.5$ m^{-1} and (b) $\phi = 2$ m^{-1}, corresponding to *effective ranges* of 6 m and 0.5 m. (The effective ranges are distances where the spatial correlation falls to 0.05, approximately equal to $3/\phi$, see [33, p. 26].)

We now study the mean-square error (MSE) performances of the proposed methods. Our performance metrics are the *average* MSEs of mean-signal estimators $\widehat{\alpha}$ and random-field estimators $\widehat{\boldsymbol{y}}_{1:N}$, computed using $L = 5000$ independent trials:

$$\text{MSE}(\widehat{\alpha}) = \text{E}\left[(\widehat{\alpha} - \alpha)^2\right] \approx \frac{1}{L} \sum_{l=1}^{L} \left(\widehat{\alpha}\big|_{l\text{th trial}} - \alpha\big|_{l\text{th trial}}\right)^2 \tag{5.4a}$$

$$\text{MSE}(\widehat{\boldsymbol{y}}_{1:N}) = \frac{1}{N}\text{E}\left[\|\widehat{\boldsymbol{y}}_{1:N} - \boldsymbol{y}_{1:N}\|^2\right] \approx \frac{1}{NL} \sum_{l=1}^{L} \left\|\widehat{\boldsymbol{y}}_{1:N}\big|_{l\text{th trial}} - \boldsymbol{y}_{1:N}\big|_{l\text{th trial}}\right\|^2 \tag{5.4b}$$

where *averaging* is performed over random realizations of the mean-field signal α [simulated from (2.3) using (5.3b)] and random realizations of the measurement locations $\boldsymbol{x}_n$, $n = 1, 2, \ldots, N$. In Fig. 3, we show the average MSEs (5.4a) for the following estimates of α:

(i)   MC estimates $\alpha_{(p)}$ in (3.9a) computed using the adaptive quantization and feedback schemes in Section II-A, with $T = 20$ draws;

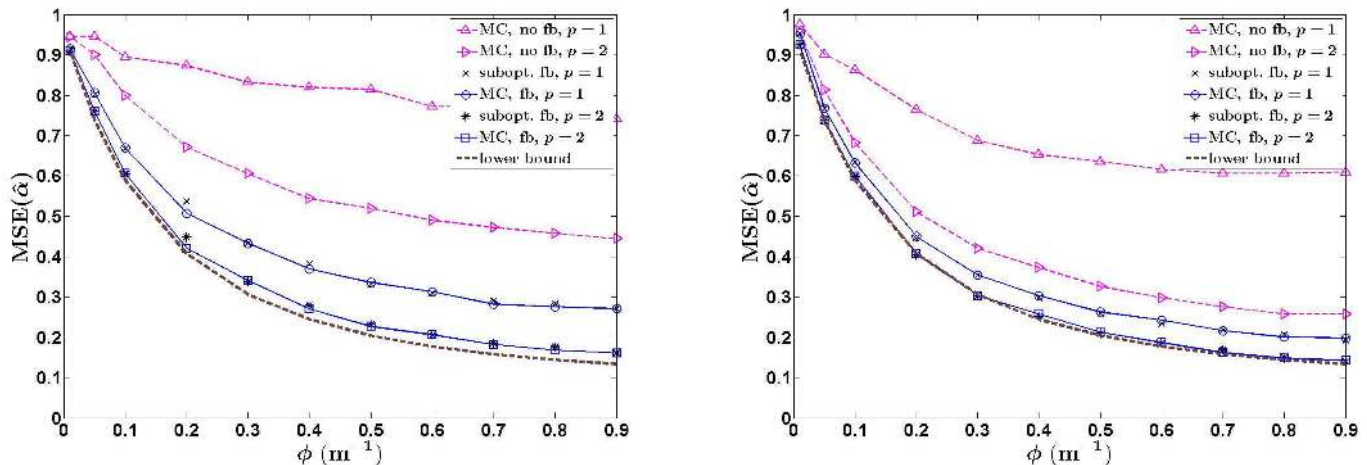(ii)   *no-feedback* MC estimates in Section II-B, with multiple transmissions and $T = 20$ draws;

Fig. 3.   Constant mean-signal scenario: Average MSEs of various estimators of α as functions of ϕ, for (a) $K = 2$ and (b) $K = 3$ quantization intervals and $p \in \{1, 2\}$ rounds.

    (iii)   suboptimal estimates $\alpha^{(0)}_{(p)}$ in (4.1c) computed using the adaptive quantization and feedback schemes in Section IV;

    (iv)   estimates $\mathrm{E}[\alpha\,|\,\boldsymbol{v}]$ for the case where analog (unquantized) observations $\boldsymbol{v}$ are available at the fusion center (used as a *benchmark*); the *average* MSE of this estimator is

$$\mathrm{E}[\Omega] \approx \frac{1}{L} \sum_{l=1}^{L} \Omega\big|_{l\text{th trial}} \tag{5.5}$$

where *averaging* is performed over the random realizations of Ψ that vary due to the random node placement in each trial, see also (3.2d).

Clearly, (5.5) is a *lower bound* on the average MSEs achievable by estimators of α that are based on quantized observations; hence, in Fig. 3, we refer to (5.5) as lower bound. In this example, the suboptimal estimates $\alpha^{(0)}_{(p)}$ achieve approximately the same performance as the corresponding MC estimates $\alpha_{(p)}$. When $K = 3$ quantization intervals are employed, the proposed estimators perform remarkably well, effectively attaining the lower bound (5.5) after $p = 2$ rounds, see Fig. 3 (b).

As ϕ decreases to zero (corresponding to highly correlated random field), this bound converges to the following constant:

$$\mathrm{E}[\Omega]\big|_{\phi \searrow 0} = (\Gamma_\alpha^{-1} + \psi^{-2})^{-1} \approx 0.96 \tag{5.6a}$$

see also Fig. 3. In this case, the measurements at the nodes $n = 1, 2, \ldots, N$ are redundant, as they are approximately equal to the fusion center's observation y_{N+1}; since all information is effectively captured by a single observations *and* our prior on α is diffuse, the constant in (5.6a) is approximately equal to the small-scale random-field variance $\psi^2 = 1$ in (5.2b).
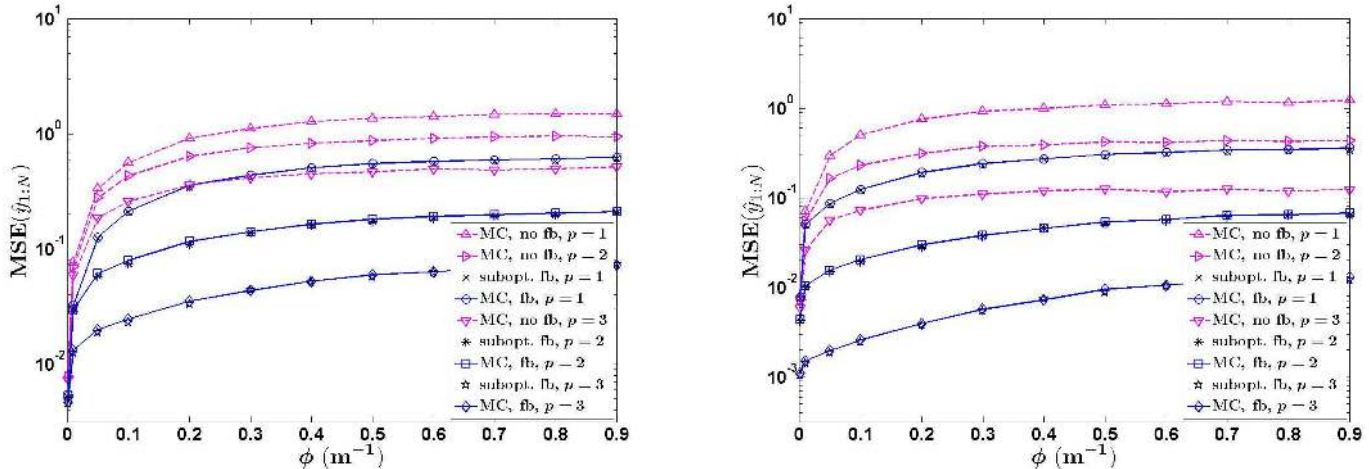
Fig. 4. Constant mean-signal scenario: Average MSEs of various estimators of $\boldsymbol{y}_{1:N}$ as functions of ϕ, for (a) $K = 2$ and (b) $K = 3$ quantization intervals and $p \in \{1, 2, 3\}$ rounds.

For large ϕ, corresponding to *white* small-scale random-field variations (i.e. Ψ proportional to the identity matrix, see (5.1)), the lower bound (5.5) converges to the following constant:

$$\mathrm{E}\left[\Omega\right]\big|_{\phi \nearrow +\infty} = \left(\Gamma_\alpha^{-1} + \frac{N+1}{\psi^2}\right)^{-1} \approx 0.09 \tag{5.6b}$$

see also Fig. 3. Since our prior on α is diffuse, the term Γ_α^{-1} in (5.6b) can be neglected; consequently, this constant is approximately $N + 1 = 11$ times smaller than ψ^2. In this case, each node's measurement contributes equal amount of information to the estimation of α. Under the constant mean-signal scenario with white small-scale random-field variations, the node indices n become *irrelevant* for estimating α; instead, the numbers of nodes reporting $k_{n,(p)} = 0, 1, 2, \ldots, K - 1$ are sufficient. We can exploit this fact to simplify communication from the nodes to the fusion center; however, this scenario has limited applicability.

Fig. 4 shows average MSEs (5.4b) of various random-field estimators as functions of ϕ, for (a) $K = 2$ and (b) $K = 3$ quantization intervals. The suboptimal estimates $\boldsymbol{y}_{1:N,(p)}^{(0)}$ in (4.1a) *match* the performance of the corresponding MC estimates $\overline{\boldsymbol{y}}_{1:N,(p)}$ in (3.8a) (computed using $T = 20$ samples). The feedback MC and suboptimal random-field estimators achieve remarkably good average MSE performances after $p = 3$ rounds.

The performances of all methods improve uniformly as we increase the number of quantization levels from $K = 2$ to $K = 3$: compare parts (a) and (b) of Figs. 3 and 4. The feedback estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{y}_{1:N}$ *significantly outperform* their no-feedback counterparts; already in Round 1, the feedback methods achieve smaller average MSEs than the no-feedback scheme after two transmissions. Therefore, feedback facilitates energy-efficient node transmissions, which is quantified in the following example.
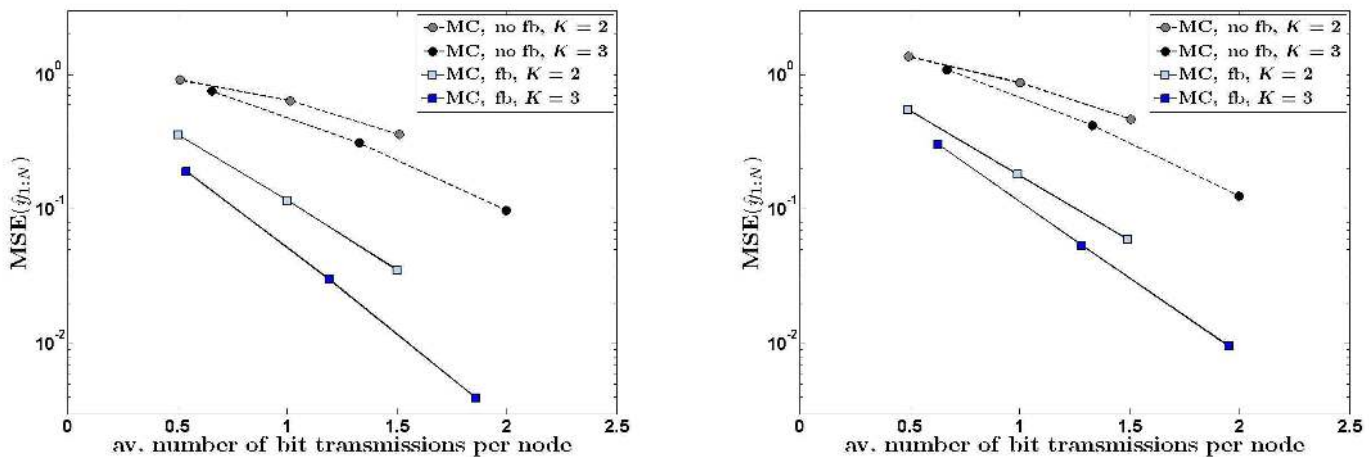
Fig. 5. Constant mean-signal scenario: Average MSEs of MC random-field estimators as functions of the average number of bit transmissions per node, for (a) $\phi = 0.2$ m^{-1} and (b) $\phi = 0.5$ m^{-1}.

1) Average MSE of random-field estimation versus average number of bit transmissions per node: In each transmission from the nodes, we employ censoring, described below for $K \in \{2, 3\}$.

$K = 2$, Round p: If $k_{n,(p)} = 1$, node n transmits a single bit; otherwise (i.e. if $k_{n,(p)} = 0$), node n stays silent.

$K = 3$, Round p: If $k_{n,(p)} \in \{0, 2\}$, node n transmits a single bit, e.g. -1 if $k_{n,(p)} = 0$ and $+1$ if $k_{n,(p)} = 2$; otherwise (i.e. if $k_{n,(p)} = 1$), node n stays silent.

Fig. 5 shows average random-field MSEs (5.4b) for various (feedback and no-feedback) MC random-field estimators as functions of the average number of bit transmissions per node for (a) $\phi = 0.2$ m^{-1} and (b) $\phi = 0.5$ m^{-1}. Here, both the average number of bit transmission per node and average MSEs (5.4b) are estimated using the $L = 5000$ trials. As the spatial correlation decreases (i.e. ϕ increases), the average MSEs increase, compare Fig. 5 (a) with Fig. 5 (b). Observe the nearly perfect exponential decay of the average MSEs as functions of the average number of transmitted bits. In this example, fusion-center feedback brings *an order of magnitude improvement* compared with the no-feedback approach. About two bit transmissions per node (on average) are sufficient to estimate the random field remarkably well.

2) Random-field prediction: Consider now $N = 10$ nodes placed over a 10 m $\times 10$ m region of interest, at locations depicted by crosses, and a fusion center placed at the center of this region, marked with a circle, see Fig. 6. In this example, the nodes quantize their observations using $K = 3$ intervals. The random field was generated on a 100×100 uniform grid covering the region of interest, using $\phi = 0.2$ m^{-1} and $\phi = 0.5$ m^{-1} and estimated and predicted at these locations using the MC approach in Sections III-A and
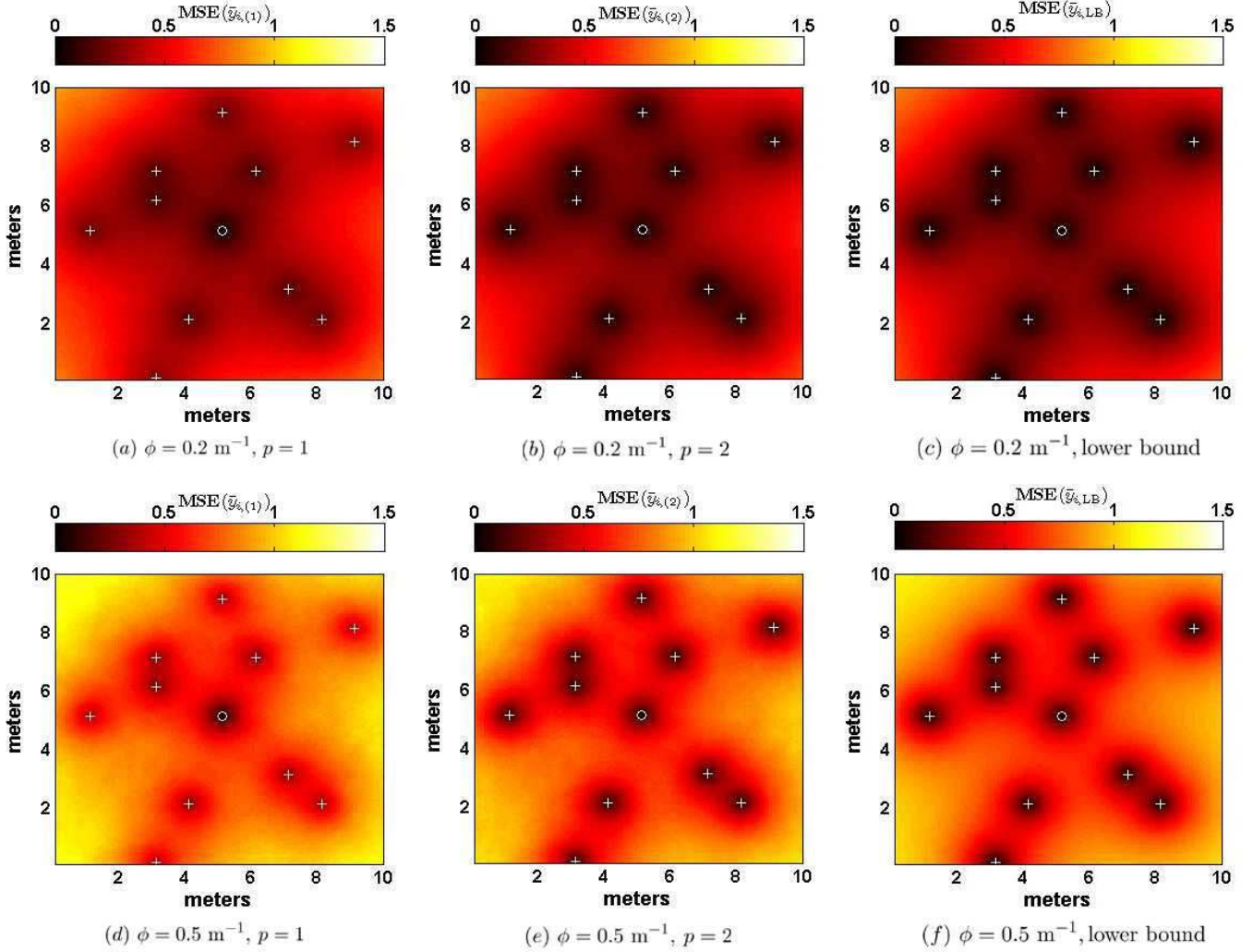
Fig. 6. Constant mean-signal scenario: average MSEs of the feedback MC random-field predictors for (a) $\phi = 0.2$ m^{-1} and $p = 1$ round, (b) $\phi = 0.2$ m^{-1} and $p = 2$ rounds, (c) lower bound for $\phi = 0.2$ m^{-1}, (d) $\phi = 0.5$ m^{-1} and $p = 1$, (e) $\phi = 0.5$ m^{-1} and $p = 2$, and (f) lower bound for $\phi = 0.5$ m^{-1}, using $K = 3$ quantization intervals.

III-B with $T = 20$ samples. Hence, we predict the random field at $M = 100^2 - N - 1 = 9989$ locations:

$$i \in \{N + 2, N + 3, \ldots, N + M + 1\}. \tag{5.7a}$$

In this example, the suboptimal scheme yields estimates $\boldsymbol{y}_{1:N,(p)}^{(0)}$ and predictions $y_{i,(p)}^{(0)}$ that are almost identical to those of the MC approach; hence, we present only the MC results $\overline{\boldsymbol{y}}_{1:N,(p)}$ and $\overline{y}_{i,(p)}$ in (3.8a) and (3.13a). Here, our performance metric is the *average* MSE of the MC random-field predictor $\overline{y}_{i,(p)} = \mathrm{E}\left[y_i \mid y_{N+1}, \mathcal{R}_{(p)}\right]$ at locations with indices i in (5.7a):

$$\mathrm{MSE}(\overline{y}_{i,(p)}) = \mathrm{E}\left[(\overline{y}_{i,(p)} - y_i)^2\right] \approx \frac{1}{L} \sum_{l=1}^{L} \left(\overline{y}_{i,(p)}\big|_{l\text{th trial}} - y_i\big|_{l\text{th trial}}\right)^2 \tag{5.7b}$$

computed using $L = 5000$ independent trials, where *averaging* is performed over random realizations of the mean-field signal α [simulated from (2.3) using (5.3b)]. Note that the node locations are *fixed* in this

example, i.e. they *do not vary* from one trial to another. Parts (a) and (b) in the upper row of Fig. 6 show performances of our MC random-field predictors and estimators after $p = 1$ and $p = 2$ feedback rounds (respectively) for $\phi = 0.2$ m^{-1} whereas parts (d) and (e) in the lower row show the corresponding performances for $\phi = 0.5$ m^{-1}. As expected, the random-field prediction performance is better for fields with stronger spatial correlations (smaller ϕ).

Parts (c) and (f) in Fig. 6 show the average prediction MSEs (3.12b) for the case where analog observations $\boldsymbol{v}$ are available at the fusion center and $\phi = 0.2$ m^{-1} and 0.5 m^{-1}, respectively. Clearly, (3.12b) is a *lower bound* on the average MSE achievable by our random-field predictor at location $\boldsymbol{x}_i$. In Fig. 6, average prediction MSE performances close to the lower bounds (3.12b) are achieved after few feedback rounds.

B. Linear Trend Surface

Consider now a linear trend-surface model for the large-scale signal component, with $\boldsymbol{h}_n = [1 \ \boldsymbol{x}_n^T]^T$:

$$H = [\mathbf{1}_{N+1}, \ [\boldsymbol{x}_1, \boldsymbol{x}_2 \cdots \boldsymbol{x}_{N+1}]^T]. \tag{5.8a}$$

and the 3×1 regression-coefficient vector $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \alpha_3]^T$, where α_1 and α_2, α_3 quantify the mean and slope of the trend surface, respectively. A realization of this random field for $\boldsymbol{\alpha} = [0, 2, -1]$ and $\phi = 0.5$ m^{-1} is given in Fig. 14 (a). Denote by $\mathrm{diag}\{\gamma_1, \gamma_2, \gamma_3\}$ a 3×3 diagonal matrix with diagonal elements equal to γ_1, γ_2, and γ_3. We choose a diffuse prior pdf for $\boldsymbol{\alpha}$, with [see (2.3)]

$$\boldsymbol{\mu_\alpha} = \mathbf{0}_{3\times 1}, \quad \Gamma_\alpha = \mathrm{diag}\{25, 1, 1\} \tag{5.8b}$$

implying that the expected marginal variance of the measurements y_n is [see also (3.3c)]:

$$\mathrm{E}\left[\mathrm{var}(y_n)\right] = \mathrm{E}\left[\sigma_n^2\right] = \mathrm{E}\left[\Sigma_{n,n}\right] \approx 43 \tag{5.8c}$$

which roughly quantifies the dynamic range of measurements at the nodes. Here, the expectation is taken over the random realizations of Ψ that vary due to random node placement in each trial. Note that the fusion-center's measurement y_{N+1} *does not* provide information about the slope of the trend surface, making this scenario more challenging than that for constant mean signal in Section V-A.

Our performance metrics are the average random-field estimation MSEs in (5.4b) and average MSE matrices for regression-vector estimators $\widehat{\boldsymbol{\alpha}}$

$$\mathrm{MSE}(\widehat{\boldsymbol{\alpha}}) = \mathrm{E}\left[(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha})^T\right]$$
$$\approx \frac{1}{L}\sum_{l=1}^{L}\left(\widehat{\boldsymbol{\alpha}}\big|_{l\text{th trial}} - \boldsymbol{\alpha}\big|_{l\text{th trial}}\right)\left(\widehat{\boldsymbol{\alpha}}\big|_{l\text{th trial}} - \boldsymbol{\alpha}\big|_{l\text{th trial}}\right)^T$$

computed using $L = 5000$ independent trials, where averaging is performed over random node locations $\boldsymbol{x}_n$. In Figs. 7 and 8, we show performances of the following estimators of $\boldsymbol{\alpha}$ as functions of ϕ, using $K = 2$ and $K = 3$ quantization intervals (respectively):

(i)   MC estimates $\boldsymbol{\alpha}_{(p)}$ in (3.9a) computed using $T = 50$ draws;

(ii)   no-feedback MC estimates with multiple transmissions, computed using $T = 50$ draws;

(iii)   suboptimal estimates $\boldsymbol{\alpha}_{(p)}^{(0)}$;

(iv)   estimates $\mathrm{E}\left[\boldsymbol{\alpha} \,|\, \boldsymbol{v}\right]$ for the case where analog observations $\boldsymbol{v}$ are available at the fusion center (used as a *benchmark*); the average MSE of this estimator is given in (5.5).

In particular, for an estimator $\widehat{\boldsymbol{\alpha}} = [\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\alpha}_3]^T$, we show average MSEs for the trend-surface mean and slope estimates $\widehat{\alpha}_1$ and $\widehat{\alpha}_2, \widehat{\alpha}_3$:

$$\mathrm{MSE}(\widehat{\alpha}_1) = [\mathrm{MSE}(\widehat{\boldsymbol{\alpha}})]_{1,1}$$

$$\tfrac{1}{2}[\mathrm{MSE}(\widehat{\alpha}_2) + \mathrm{MSE}(\widehat{\alpha}_3)] = \tfrac{1}{2}\{[\mathrm{MSE}(\widehat{\boldsymbol{\alpha}})]_{2,2} + [\mathrm{MSE}(\widehat{\boldsymbol{\alpha}})]_{3,3}\}$$

as well as the corresponding *lower bounds* $\mathrm{E}\left[\Omega_{1,1}\right]$ and $\mathrm{E}\left[\tfrac{1}{2}\left(\Omega_{2,2} + \Omega_{3,3}\right)\right]$ on the average MSEs achievable by estimators of $\boldsymbol{\alpha}$ that are based on quantized observations. As ϕ decreases to zero, these bounds converge to $\mathrm{E}\left[\Omega_{1,1}\right]\big|_{\phi\searrow 0} \approx 0.96$ [where we used (5.2b) and (5.3b)] and $\mathrm{E}\left[\tfrac{1}{2}\left(\Omega_{2,2} + \Omega_{3,3}\right)\right]\big|_{\phi\searrow 0} = 0$, see Figs. 7 and 8. Here, the small-scale random-field variations approximately satisfy $e_1 = e_2 = \ldots = e_{N+1}$, implying that the trend-surface slope regression coefficients α_1 and α_2 can be determined very accurately. For large ϕ, (5.5) simplifies to

$$\mathrm{E}\left[\Omega\right]\Big|_{\phi\nearrow+\infty} = \mathrm{E}\left\{\left(\Gamma_{\boldsymbol{\alpha}}^{-1} + \frac{1}{\psi^2}\left[\begin{array}{cc} N+1 & \sum_{n=1}^{N+1}\boldsymbol{x}_n^T \\ \sum_{n=1}^{N+1}\boldsymbol{x}_n & (\sum_{n=1}^{N+1}\boldsymbol{x}_n\,\boldsymbol{x}_n^T) \end{array}\right]\right)^{-1}\right\} \approx \left[\begin{array}{ccc} 0.119 & 0.007 & 0.005 \\ 0.007 & 0.016 & 0 \\ 0.005 & 0 & 0.016 \end{array}\right]$$

where we used (5.2a), (5.2b), and (5.3b), see Figs. 7 and 8.

Since fusion-center's measurement y_{N+1} *does not* provide information about the trend-surface slope regression coefficients α_2 and α_3, we must employ feedback to estimate α_2 and α_3 well. Here, the proposed feedback MC estimators outperform their suboptimal counterparts in estimating these parameters. The lower bounds are effectively attained for $K = 3$ quantization levels and $p = 3$ rounds, see Fig. 8.

Fig. 9 shows average MSEs (5.4b) of various random-field estimators as functions of ϕ, using (a) $K = 2$ and (b) $K = 3$ quantization intervals. Again, the MC scheme (computed using $T = 50$ samples) outperforms the suboptimal scheme.
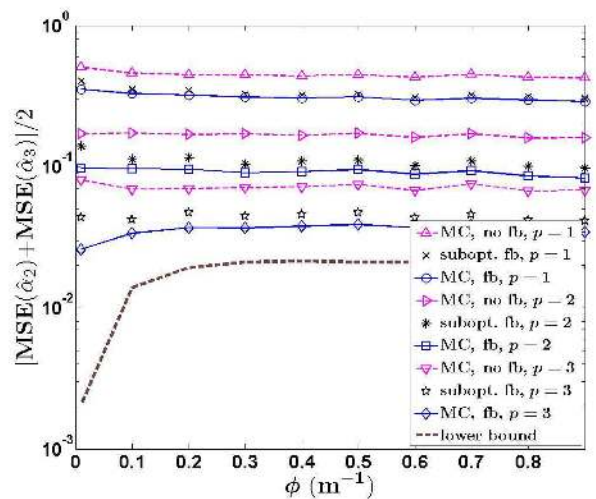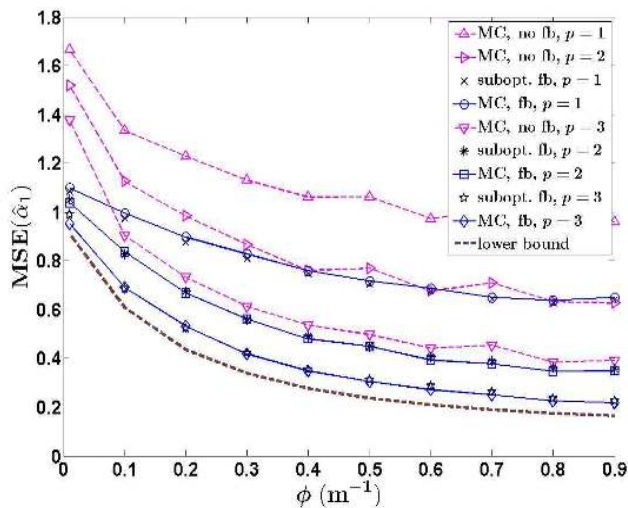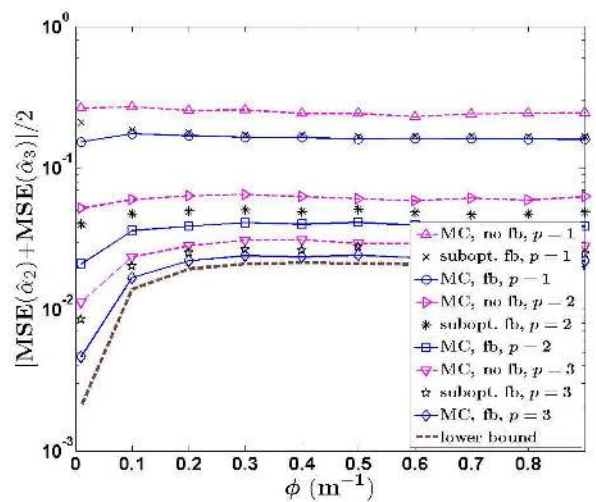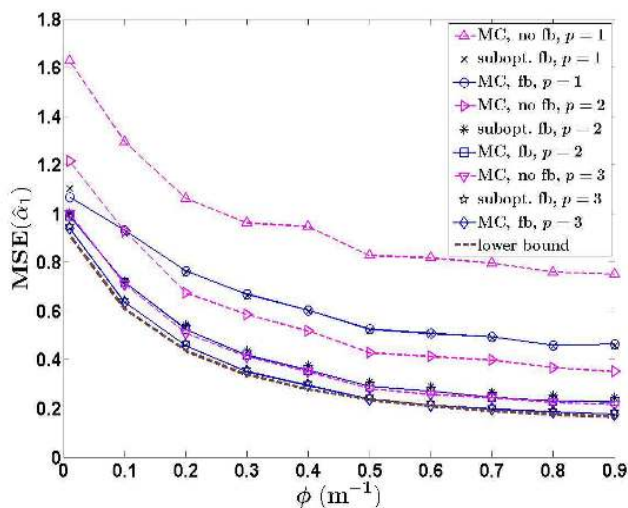
Fig. 7. Linear trend-surface scenario: (a) $\text{MSE}(\widehat{\alpha}_1)$ and (b) $\frac{1}{2}\left[\text{MSE}(\widehat{\alpha}_2) + \text{MSE}(\widehat{\alpha}_3)\right]$ as functions of ϕ, for $K = 2$ quantization intervals.

Fig. 8. Linear trend-surface scenario: (a) $\text{MSE}(\widehat{\alpha}_1)$ and (b) $\frac{1}{2}\left[\text{MSE}(\widehat{\alpha}_2) + \text{MSE}(\widehat{\alpha}_3)\right]$ as functions of ϕ, for $K = 3$ quantization intervals.

The performances of all methods improve uniformly as we increase the number of quantization levels from $K = 2$ to $K = 3$: compare Fig. 7 with Fig. 8 and Fig. 9 (a) with Fig. 9 (b). As expected, the feedback estimators of $\boldsymbol{\alpha}$ and $\boldsymbol{y}_{1:N}$ outperform their no-feedback counterparts.

1) Average MSE of random-field estimation versus average number of bit transmissions per node: We employ censoring in each transmission from the nodes, as described in Section V-A.1. Fig. 10 shows average random-field estimation MSEs (5.4b) for various MC estimators (computed using $T = 50$ samples) as functions of the average number of bit transmissions per node for (a) $\phi = 0.2$ m^{-1} and (b) $\phi = 0.5$ m^{-1}. We again observe the nearly perfect exponential decay of the average MSEs as functions of the average number of transmitted bits. As the spatial correlation decreases (i.e. ϕ increases), the average MSEs increase
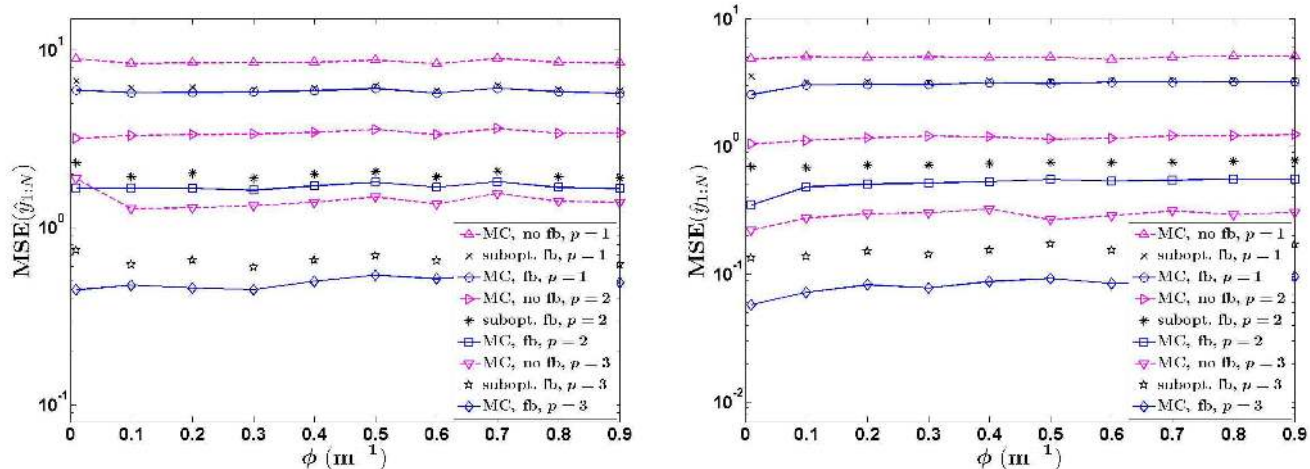
Fig. 9. Linear trend-surface scenario: Average MSEs of the proposed estimators of $\boldsymbol{y}_{1:N}$ as functions of ϕ, for (a) $K = 2$ and (b) $K = 3$ quantization intervals and $p \in \{1, 2, 3\}$ rounds.
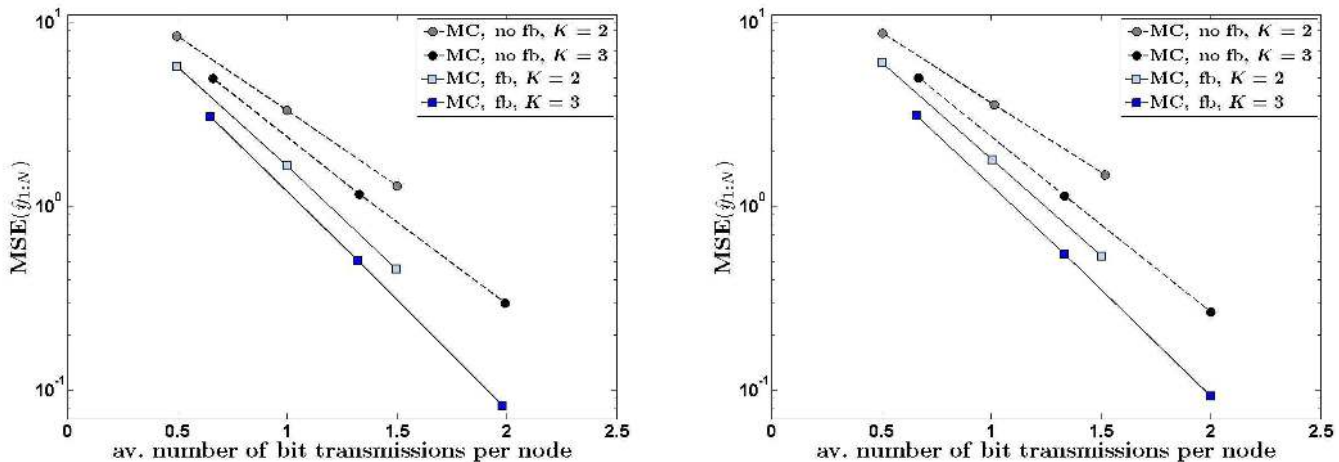

Fig. 10. Linear trend-surface scenario: Average MSEs of MC random-field estimators as functions of the average number of bit transmissions per node, for (a) $\phi = 0.2$ m^{-1} and (b) $\phi = 0.5$ m^{-1}.

only slightly, compare Fig. 10 (a) with Fig. 10 (b). For the same average number of transmitted bits, the average MSEs of the feedback random-field estimation approach are up to three times smaller than those of the no-feedback scheme. In this example, feedback brings less improvement compared with the constant-mean case in Section V-A.1.

2) Average MSE of regression-coefficient and random-field estimators as functions of the small-scale random-field variance ψ^2: Throughout this section, we set the correlation-strength parameter ϕ to

$$\phi = 0.5 \text{ m}^{-1}. \tag{5.9}$$

Figs. 11 and 12 show MSE$(\widehat{\alpha}_1)$ and $\frac{1}{2}\left[\text{MSE}(\widehat{\alpha}_2) + \text{MSE}(\widehat{\alpha}_3)\right]$ for various estimators $\widehat{\boldsymbol{\alpha}}$ as functions of ψ^2, using $K = 2$ and $K = 3$ quantization intervals, respectively. Assuming that $\psi^2 > 0.1$, the lower bounds
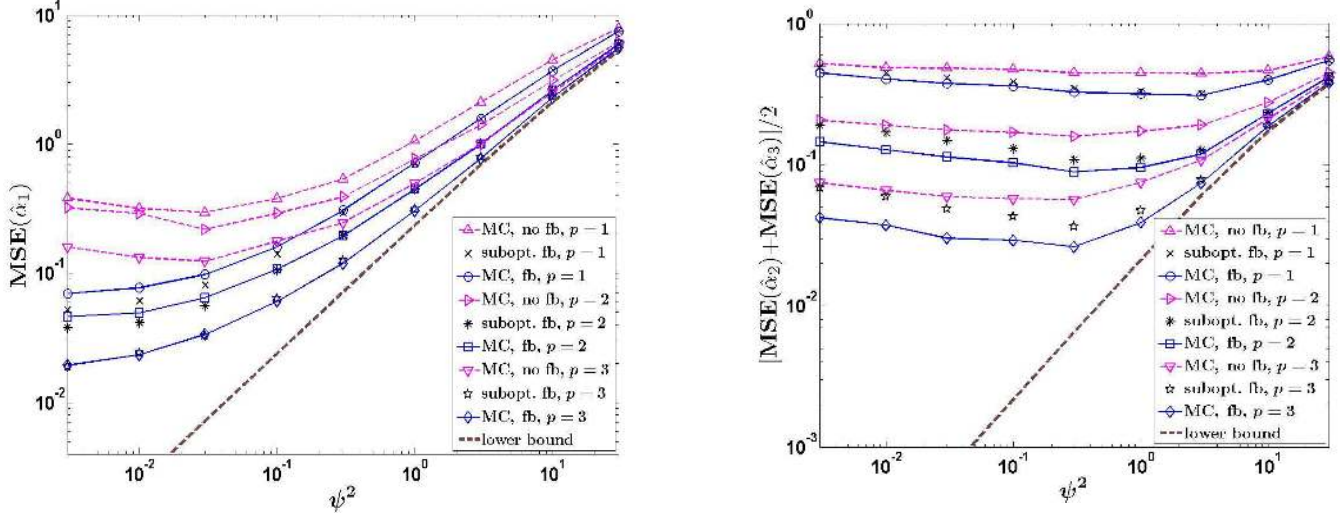
Fig. 11. Linear trend-surface scenario: (a) $\mathrm{MSE}(\widehat{\alpha}_1)$ and (b) $\frac{1}{2}\left[\mathrm{MSE}(\widehat{\alpha}_2)+\mathrm{MSE}(\widehat{\alpha}_3)\right]$ as functions of ψ^2, for $K=2$ quantization intervals.
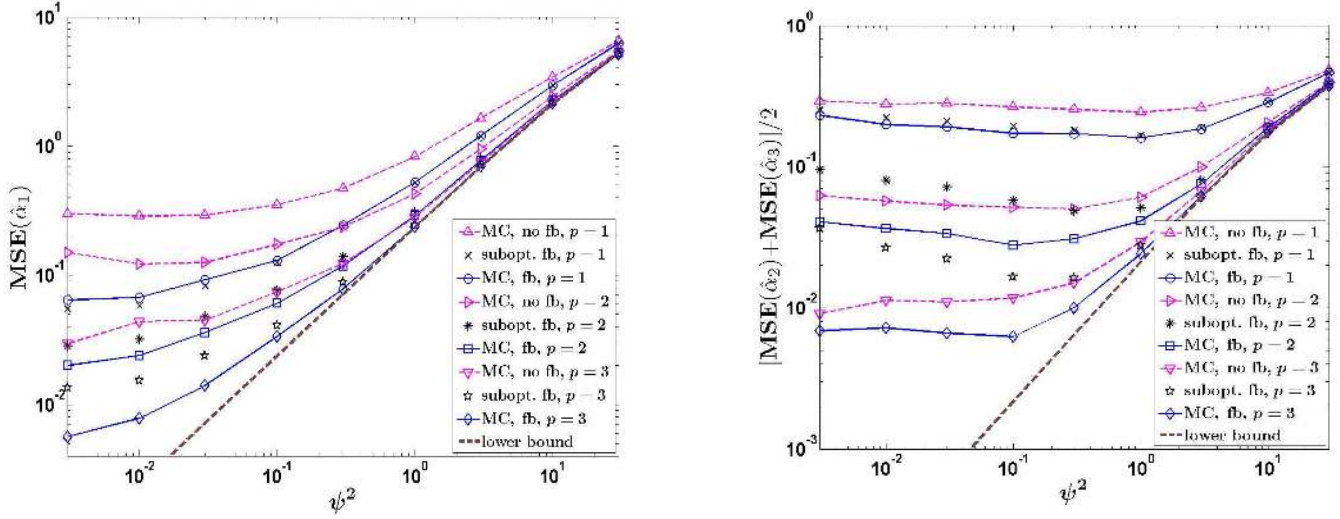
Fig. 12. Linear trend-surface scenario: (a) $\mathrm{MSE}(\widehat{\alpha}_1)$ and (b) $\frac{1}{2}\left[\mathrm{MSE}(\widehat{\alpha}_2)+\mathrm{MSE}(\widehat{\alpha}_3)\right]$ as functions of ψ^2, for $K=3$ quantization intervals.

are approximately attained for $K=3$ quantization levels and $p=3$ feedback rounds, see Fig. 12. Fig. 13 shows average MSEs (5.4b) of various random-field estimators as functions of ψ^2, using (a) $K=2$ and (b) $K=3$ quantization intervals. As expected, the performances of all methods improve uniformly as we increase the number of quantization levels from $K=2$ to $K=3$.

As ψ^2 increases, the role of the prior pdf *grows* and, for sufficiently large ψ^2, the lower bound (5.5) simplifies to the prior covariance matrix:

$$\mathrm{E}\left[\Omega\right]\big|_{\psi^2\nearrow+\infty}=\Gamma_{\boldsymbol{\alpha}} \tag{5.10}$$

meaning that we rely solely on the prior information about $\boldsymbol{\alpha}$. Therefore, in this case, *no feedback is needed*
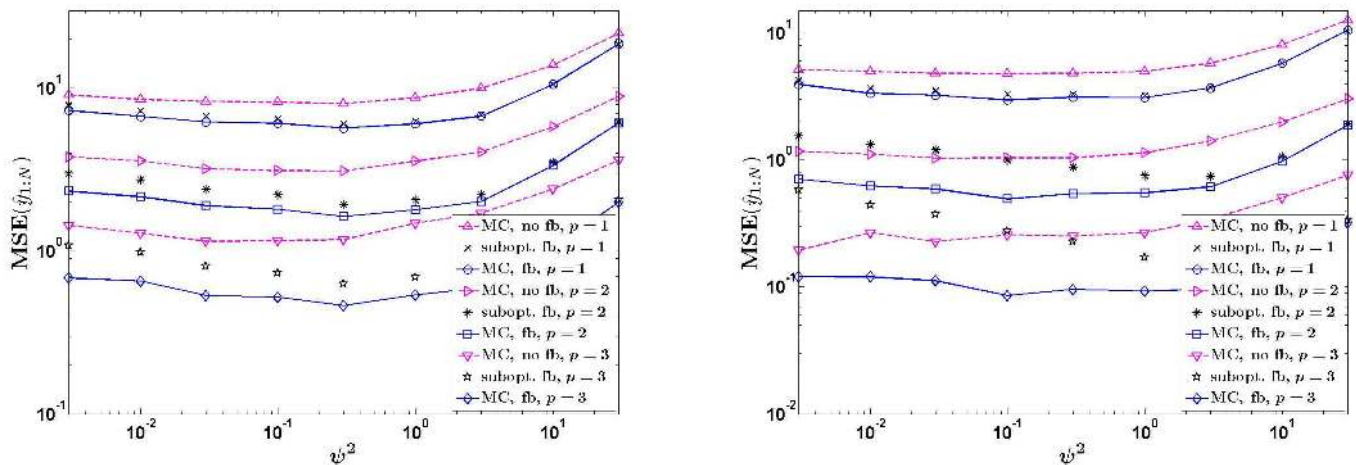
Fig. 13. Linear trend-surface scenario: Average MSEs of the proposed estimators of $\boldsymbol{y}_{1:N}$ as functions of ψ^2 for (a) $K = 2$ and (b) $K = 3$ quantization intervals and $p \in \{1, 2, 3\}$ rounds.

regarding $\boldsymbol{\alpha}$ and our feedback and no-feedback approaches become equivalent, see Figs. 11–13.

Observe a *dithering-like effect* exhibited (to various degrees) by the feedback methods in Fig. 13; this effect also shows in slope regression-coefficient estimation, see Fig. 11 (b) and Fig. 12 (b). In particular, for ψ^2 *sufficiently small*, the feedback random-field and slope-parameter estimators *improve* with the increase of ψ^2. For $p = 3$ rounds and $K = 3$ quantization levels, the suboptimal field and slope-parameter estimates perform poorly when ψ^2 is small, see Figs. 13 (b) and 12. This poor performance is likely due to the crude posterior covariance estimates $C_{(p-1)}^{(0)}$ in (4.1d) that the suboptimal approach employs in Rounds $p = 2, 3, \ldots$. Indeed, in Round 1 where the suboptimal method utilizes $C_{(0)}^{(0)} = C_{(0)}$ in (2.4c) [the same as that used in Round 1 of the feedback MC method], the suboptimal method matches the performance of the feedback MC method.

To illustrate the impact of fusion-center feedback on field prediction, we now apply the MC feedback scheme to a random-field realization generated from the linear-trend and correlation models in (5.3a) and (5.1) with $\boldsymbol{\alpha} = [0, 2, -1]$ and $\phi = 0.5$ m^{-1}, shown in Fig. 14 (a). The nodes sample this random field at locations depicted by crosses. In this example, the nodes quantize their observations using $K = 3$ intervals. We applied the MC random-field estimators (3.8a) and predictors (3.13a) using $T = 50$ draws. The obtained results after $p = 1, 2,$ and 3 feedback rounds are shown in parts (b)–(d) of Fig. 14; part (e) shows the lower bounds (3.12b) on average MSEs achievable by our predictors, whereas parts (f)–(h) show MC posterior-predictive variances (3.13b) after $p = 1, 2,$ and 3 rounds, respectively.
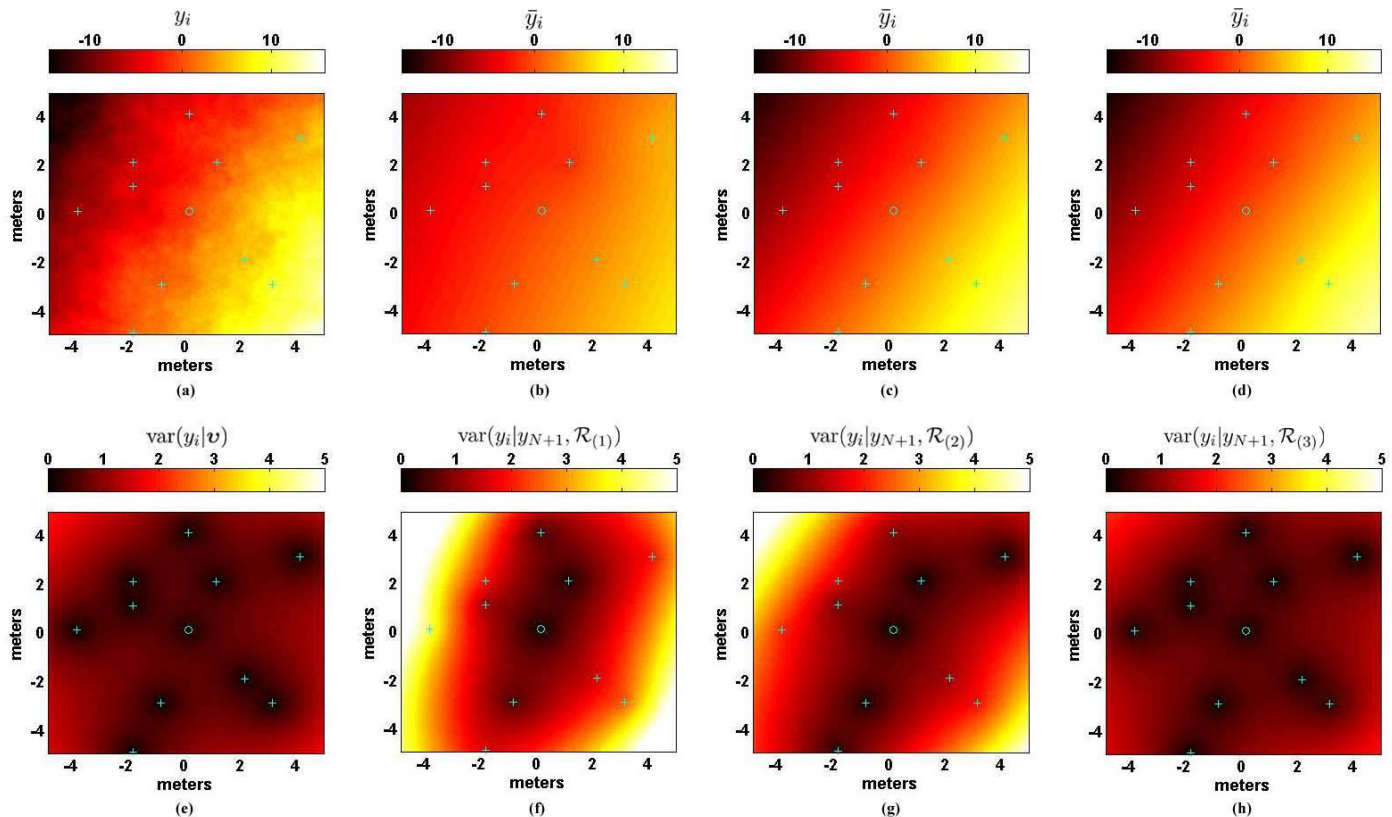
Fig. 14. Linear trend-surface scenario with $K = 3$ quantization intervals: (a) a random-field realization; feedback MC random-field estimates and predictions after (b) $p = 1$, (c) $p = 2$, and (d) $p = 3$ rounds; (e) lower bounds on random-field prediction MSEs, and MC posterior-predictive variances after (f) $p = 1$, (g) $p = 2$, and (h) $p = 3$ rounds.

VI. CONCLUDING REMARKS

We developed a Bayesian framework for adaptive quantization, fusion-center feedback, and estimation of a spatially correlated Gaussian random field and its parameters.

In this paper, we *did not* take into account the energy consumption at the nodes within the region of interest due to "listening," i.e. receiving fusion-center feedback. Instead, we focused on utilizing this feedback to reduce the node transmission energy. Further research will include:

- developing *adaptive* no-feedback quantization schemes [44];

- accounting for the listening cost and applying *quantized* fusion-center feedback;

- analyzing the impact of communication errors between the nodes and the fusion center;[4]

- analyzing the proposed feedback schemes and analytically predicting the results in Figs. 5 and 10;

- extending the proposed framework to random-field tracking over time and to nonlinear models;

- extending our measurement model to correlated non-Gaussian random fields by applying copulas [43].

[4]Interestingly, some communication errors may be detected during the adaptive quantization process. A communication error has occurred if (2.9c) yields an empty set or multiple intervals (instead of a single interval).

APPENDIX

To simplify notation, in this appendix we omit the round subscript "$_{(p)}$" and use $\boldsymbol{y}_{1:N}^{(t)}$, $\boldsymbol{\alpha}^{(t)}$, $\mathcal{R}_n$, $\mathcal{R}$, and $\boldsymbol{k} = [k_1, k_2, \ldots, k_N]$ in place of $\boldsymbol{y}_{1:N,(p)}^{(t)}$, $\boldsymbol{\alpha}_{(p)}^{(t)}$, $\mathcal{R}_{n,(p)}$, $\mathcal{R}_{(p)}$, and $\boldsymbol{k}_{(p)} = [k_{1,(p)}, k_{2,(p)}, \ldots, k_{N,(p)}]^T$.

Let us explicitly define the endpoints of the interval $\mathcal{R}_n$:

$$\mathcal{R}_n = \big(y_{n,\text{MIN}}, y_{n,\text{MAX}}\big).$$

APPENDIX A. SIMULATING $x_1^{(t)}, x_2^{(t)}, \ldots, x_N^{(t)}$ IN (3.6)

We now show how to simulate $x_1^{(t)}, x_2^{(t)}, \ldots, x_N^{(t)}$ in (3.6a) (Step 1), (3.6b) (Step 2), ..., (3.6c) (Step N). Note that [see (3.5)]

$$l_{nn} > 0, \quad n = 1, 2, \ldots, N \tag{A.1}$$

because S is a positive definite matrix.

We first unify the tasks in Steps n, $n = 2, 3, \ldots, N$ and Step 1. Then, we outline a *rejection-sampling approach* to their implementation.

Steps 2,...,N. For $2 \leq n \leq N$, we draw samples from a truncated univariate Gaussian distribution with density proportional to (see also (A.1))

$$\mathcal{N}(x_n \,;\, [\mathcal{L}^{-1}\boldsymbol{m}]_n, 1) \cdot i_{\mathcal{R}_n}\big(l_{n1} x_1^{(t)} + \cdots l_{n(n-1)} x_{n-1}^{(t)} + l_{nn} x_n\big) \tag{A.2}$$

where $x_1^{(t)}, x_2^{(t)}, \ldots x_{n-1}^{(t)}$ are fixed values, treated as constants. Equivalently, we sample z_n from a truncated standard normal pdf with density proportional to [see also (A.1)]

$$w_n(z_n) = \mathcal{N}(z_n \,;\, 0, 1) \cdot i_{(z_{n,\text{MIN}}^{(t)}, z_{n,\text{MAX}}^{(t)})}(z_n) \tag{A.3}$$

where

$$z_{n,\text{MIN}}^{(t)} = (y_{n,\text{MIN}} - l_{n1} x_1^{(t)} - \cdots - l_{n(n-1)} x_{n-1}^{(t)} - l_{nn} [\mathcal{L}^{-1}\boldsymbol{m}]_n)/l_{nn} \tag{A.4a}$$

$$z_{n,\text{MAX}}^{(t)} = (y_{n,\text{MAX}} - l_{n1} x_1^{(t)} - \cdots - l_{n(n-1)} x_{n-1}^{(t)} - l_{nn} [\mathcal{L}^{-1}\boldsymbol{m}]_n)/l_{nn} \tag{A.4b}$$

and then compute

$$x_n^{(t)} = [\mathcal{L}^{-1}\boldsymbol{m}]_n + z_n. \tag{A.5}$$

Step 1. We sample z_1 from univariate probability distribution with density proportional to (A.3) with $n = 1$, where

$$z_{1,\text{MIN}} = \frac{y_{1,\text{MIN}} - l_{11} [\mathcal{L}^{-1}\boldsymbol{m}]_1}{l_{11}} \tag{A.6a}$$

$$z_{1,\text{MAX}} = \frac{y_{1,\text{MAX}} - l_{11} [\mathcal{L}^{-1}\boldsymbol{m}]_1}{l_{11}} \tag{A.6b}$$

and then compute $x_1^{(t)}$ using (A.5) with $n = 1$.

Therefore, we have reduced the tasks in Steps n, $n = 1, 2, \ldots, N$ to sampling z_n, $n = 1, 2, \ldots, N$ from the truncated standard normal pdfs (A.3). To apply rejection sampling, we need a "dominating" pdf $g_n(\cdot)$, where $g_n(\cdot)$ is easy to sample from and

$$w_n(\xi) \le m_n \, g_n(\xi)$$

for all ξ and some positive constant m_n, see [41, Ch. 11.1] and [42]. We choose the dominating functions using heuristic rules similar to those in [37, Sect. 2].

APPENDIX B. COMPUTING $x_n^{(0)}$ IN SECTION IV

Based on the discussion in Appendix A, we compute $x_n^{(0)}$ in (4.1b) as follows:

$$x_n^{(0)} = [\mathcal{L}^{-1} \boldsymbol{m}]_n + z_n^{(0)} \tag{B.1a}$$

where

$$z_n^{(0)} = \frac{\int_{-\infty}^{+\infty} z \, \mathcal{N}(x \, ; \, 0, 1) \, i_{(z_{n,\mathrm{MIN}}^{(0)}, z_{n,\mathrm{MAX}}^{(0)})}(z) \, dz}{\int_{-\infty}^{+\infty} \mathcal{N}(x \, ; \, 0, 1) \, i_{(z_{n,\mathrm{MIN}}^{(0)}, z_{n,\mathrm{MAX}}^{(0)})}(z) \, dz} = \frac{1}{\sqrt{2\,\pi}} \frac{\exp[-(z_{n,\mathrm{MIN}}^{(0)})^2/2] - \exp[-(z_{n,\mathrm{MAX}}^{(0)})^2/2]}{\Phi(z_{n,\mathrm{MAX}}^{(0)}) - \Phi(z_{n,\mathrm{MIN}}^{(0)})} \tag{B.1b}$$

and, for $n = 2, 3, \ldots, N$,

$$z_{n,\mathrm{MIN}}^{(0)} = (y_{n,\mathrm{MIN}} - l_{n1} x_1^{(0)} - \cdots - l_{n(n-1)} x_{n-1}^{(0)} - l_{nn} [\mathcal{L}^{-1} \boldsymbol{m}]_n)/l_{nn} \tag{B.1c}$$

$$z_{n,\mathrm{MAX}}^{(0)} = (y_{n,\mathrm{MAX}} - l_{n1} x_1^{(0)} - \cdots - l_{n(n-1)} x_{n-1}^{(0)} - l_{nn} [\mathcal{L}^{-1} \boldsymbol{m}]_n)/l_{nn} \tag{B.1d}$$

and, for $n = 1$,

$$z_{1,\mathrm{MIN}}^{(0)} = \frac{y_{1,\mathrm{MIN}} - l_{11} [\mathcal{L}^{-1} \boldsymbol{m}]_1}{l_{11}} \tag{B.1e}$$

$$z_{1,\mathrm{MAX}}^{(0)} = \frac{y_{1,\mathrm{MAX}} - l_{11} [\mathcal{L}^{-1} \boldsymbol{m}]_1}{l_{11}} \tag{B.1f}$$

see also (A.4) and (A.6).

For $z_{n,\mathrm{MIN}}^{(0)}$ large (e.g. $z_{n,\mathrm{MIN}}^{(0)} > 7$), the above expression can be accurately computed using the expansion from [45, 26.2.13 on p. 932]:

$$z_n^{(0)} = (1 - \exp\{(z_{n,\mathrm{MIN}}^{(0)})^2/2 - (z_{n,\mathrm{MAX}}^{(0)})^2/2\}) \cdot \{q(z_{n,\mathrm{MIN}}^{(0)}) - q(z_{n,\mathrm{MAX}}^{(0)}) \exp\left[(z_{n,\mathrm{MIN}}^{(0)})^2/2 - (z_{n,\mathrm{MAX}}^{(0)})^2/2\right]\}^{-1} \tag{B.2a}$$

where

$$q(z) = \frac{1}{z} \left[1 - \frac{1}{z^2 + 2} + \frac{1}{(z^2 + 2)(z^2 + 4)} - \frac{5}{(z^2 + 2)(z^2 + 4)(z^2 + 6)} + \ldots\right]. \tag{B.2b}$$

Here, (B.2a) can be trivially modified to the case of $z_{n,\mathrm{MAX}}^{(0)}$ small (e.g. $z_{n,\mathrm{MAX}}^{(0)} < -7$), by replacing the interval $(z_{n,\mathrm{MIN}}^{(0)}, z_{n,\mathrm{MAX}}^{(0)})$ with $(-z_{n,\mathrm{MAX}}^{(0)}, -z_{n,\mathrm{MIN}}^{(0)})$.

REFERENCES

[1] Z.-Q. Luo, "Universal decentralized estimation in a bandwidth constrained sensor network," *IEEE Trans. Inform. Theory*, vol. 51, pp. 2210–2219, Jun. 2005.

[2] Z.-Q. Luo, "An isotropic universal decentralized estimation scheme for a bandwidth constrained ad hoc sensor network," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 735–744, Apr. 2005.

[3] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A.J. Goldsmith, "Power scheduling of universal decentralized estimation in sensor networks," *IEEE Trans. Signal Processing*, vol. 54, pp. 413–422, Feb. 2006.

[4] A. Krasnopeev, J.-J. Xiao, and Z.-Q. Luo, "Minimum energy decentralized estimation in a wireless sensor network with correlated sensor noises," *EURASIP J. Wireless Commun. Networking*, vol. 4, pp. 473–482, Nov. 2005.

[5] A. Ribeiro and G.B. Giannakis, "Bandwidth-constrained distributed estimation for wireless sensor networks — part II: Unknown probability density function," *IEEE Trans. Signal Processing*, vol. 54, pp. 2784–2796, Jul. 2006.

[6] M.C. Vuran, Ö.B. Akan, and I.F. Akyildiz, "Spatio-temporal correlation: Theory and applications for wireless sensor networks," *Computer Networks*, vol. 45, no. 3, pp. 245–259, 2004.

[7] A. D'Costa, V. Ramachandran, and A.M. Sayeed, "Distributed classification of Gaussian space-time sources in wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 1026–1036, Aug. 2004.

[8] M.C. Vuran and I.F. Akyildiz, "Spatial correlation-based collaborative medium access control in wireless sensor networks," *IEEE/ACM Transactions on Networking*, vol. 14, pp. 316–329, Apr. 2006.

[9] S. Pattem, B. Krishnamachari, and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *Proc. 3rd International Symposium on Information Processing in Sensor Networks,* Berkeley, CA, Apr. 2004, pp. 28–35.

[10] A.D. Sarwate and M. Gastpar, "Estimation from misaligned observations with limited feedback," in *Proc. 39th Annu. Conf. Inform. Sci. Syst.*, Baltimore, MD, Mar. 2005.

[11] S.S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," *IEEE Signal Processing Mag.*, vol. 19, pp. 51–60, Mar. 2002.

[12] M. Longo, T.D. Lookabaugh, and R.M. Gray, "Quantization for decentralized hypothesis testing under communication constraints," *IEEE Trans. Inform. Theory*, vol. 36, pp. 241–255, Mar. 1990.

[13] D. Marco, E.J. Duarte-Melo, M. Liu, and D.L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *Proc. 2nd International Symposium on Information Processing in Sensor Networks,* Palo Alto, CA, Apr. 2003, pp. 1–16.

[14] J.-F. Chamberland and V.V. Veeravalli, "Wireless sensors in distributed detection applications," *IEEE Signal Processing Mag.*, vol. 24, pp. 16–25, May 2007.

[15] Y.-W. Hong, W.-J. Huang, F.-H. Chiu, and C.-C.J. Kuo, "Cooperative communications in resource-constrained wireless networks," *IEEE Signal Processing Mag.*, vol. 24, pp. 47–57, May 2007.

[16] T. Simsek, R. Jain, and P. Varaiya, "Scalar estimation and control with noisy binary observations," *IEEE Trans. Automatic Control*, vol. 49, pp. 1598–1603, Sept. 2004.

[17] O.C. Imer and T. Başar, "Optimal estimation with limited measurements," in *Proc. 44th IEEE Conf. Decision and Control (CDC)*, Seville, Spain, Dec. 2005, pp. 1029–1034.

[18] A. Ribeiro, G.B. Giannakis, and S.I. Roumeliotis, "SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations," *IEEE Trans. Signal Processing*, vol. 4, pp. 4782–4795, Dec. 2006.

[19] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M.I. Jordan, and S.S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automatic Control*, vol. 49, pp. 1453–1464, Sept. 2004.

[20] W.U. Bajwa, J.D. Haupt, A.M. Sayeed, and R.D. Nowak, "Joint source-channel communication for distributed estimation in sensor networks," *IEEE Trans. Inform. Theory*, vol. 53, pp. 3629–3653, Oct. 2007.

[21] R. Nowak, U. Mitra, and R. Willett, "Estimating inhomogeneous fields using wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, pp. 999–1006, Aug. 2004.

[22] J. P. M. Schalkwijk and T. Kailath, "A coding scheme for additive noise channels with feedback — Part I: No bandwidth constraint," *IEEE Trans. Inform. Theory*, vol. 12, pp. 172–182, Apr. 1966.

[23] M. Gastpar, "Causal coding and feedback in Gaussian sensor networks," in *Advances in Control, Communication Networks, and Transportation Systems: In Honor of Pravin Varaiya*, E.H. Abed (Ed.), Boston, MA: Birkhäuser, 2005, ch. 10.

[24] A. Sahai and S. Mitter, "The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link — Part I: Scalar systems," *IEEE Trans. Inform. Theory*, vol. 52, pp. 3369–3395, Aug. 2006.

[25] M. Huang and S. Dey, "Dynamic quantizer design for hidden Markov state estimation via multiple sensors with fusion center feedback," *IEEE Trans. Signal Processing*, vol. 54, pp. 2887–2896, Aug. 2006.

[26] R. Srinivasan, "Distributed detection with decision feedback," *IEE Proc. Radar, Sonar Navig., Ser. F*, vol. 147, pp. 427–432, Dec. 1990.

[27] P.F. Swaszek and P. Willett, "Parley as an approach to distributed detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 31, pp. 447–457, Jan. 1995.

[28] D.A. Pados, K.W. Halford, D. Kazakos, and P. Papantoni-Kazakos, "Distributed binary hypothesis testing with feedback," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 21–42, Jan. 1995.

[29] S. Alhakeem and P.K. Varshney, "Decentralized Bayesian detection with feedback," *IEEE Trans. Syst., Man, Cybern., Pt. A*, vol. 26, pp. 503–513, Jul. 1996.

[30] M. Kisialiou and Z.-Q. Luo, "Reducing power consumption in a sensor network by information feedback," in *Proc. EUSIPCO,* Florence, Italy, Sept. 2006.

[31] R. Rajagopal, M.J. Wainwright, and P. Varaiya, "Universal quantile estimation with feedback in the communication-constrained setting," in *Proc. IEEE Int. Symp. on Information Theory*, Seattle, WA, Jul. 2006, pp. 836–840.

[32] R. Haining, *Spatial Data Analysis: Theory and Practice*, Cambridge, UK: Cambridge Univ. Press, 2003.

[33] S. Banerjee, B.P. Carlin, and A.E. Gelfand, *Hierarchical Modeling and Analysis for Spatial Data*, New York: Chapman & Hall, 2004.

[34] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: An efficient framework for modeling sensor network data," in *Proc. 3rd International Symposium on Information Processing in Sensor Networks,* Berkeley, CA, Apr. 2004, pp. 1–10.

[35] S. Appadwedula, V.V. Veeravalli, and D.L. Jones, "Decentralized detection with censoring sensors," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1362–1373, Apr. 2005.

[36] S.M. Kay, *Fundamentals of Statistical Signal Processing — Estimation Theory*, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[37] J. Geweke, "Efficient simulation from the multivariate normal and Student-t distributions subject to linear constraints," in *Computing Science and Statistics: Proc. 23rd Symposium on the Interface*, Seattle, WA, Apr. 1991, pp. 571–578.

[38] C.P. Robert, "Simulation of truncated normal variables," *Statistics and Computing*, vol. 5, pp. 121–125, Jun. 1995.

[39] J.S. Liu, *Monte Carlo Strategies in Scientific Computing*. New York: Springer-Verlag, 2001.

[40] A.T.A. Wood and G. Chan, "Simulation of stationary Gaussian processes in $[0, 1]^d$," *J. Comput. Graph. Statist.*, vol. 3, pp. 409–432, 1994.

[41] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin, *Bayesian Data Analysis*, New York: Chapman & Hall, Second ed., 2004.

[42] J. von Neumann, "Various techniques used in connection with random digits," in *John von Neumann, Collected Works,* vol. V, A.H. Taub (Ed.), New York: Pergamon, 1961, pp. 768–770.

[43] R.B. Nelsen, *An Introduction to Copulas*, New York: Springer-Verlag, 1999.

[44] A. Dogandžić and K. Qiu, "Estimating a random field in sensor networks using quantized spatially correlated data," to appear in *Proc. 42nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 2008.

[45] M. Abramowitz and I.A. Stegun (Eds.), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th printing, Dover, New York, 1972.